

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:31:32 ; Search time 5716 Seconds
(without alignments)
10491.575 Million cell updates/sec

Title: US-10-510-408-1

Perfect score: 1055
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacctgaaa 1055

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sv.*
12: gb_un.*
13: gb_vl.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055	100.0	1055	6	AX930045 Sequence
2	1055	100.0	1522	6	AX954657 Sequence
C 3	1047.4	99.3	110000	1	Continuation (28 o
C 4	1047.4	99.3	110000	1	Continuation (28 o
5	459	43.5	459	6	AX434675 Sequence
6	457	43.3	17787	1	BSY15896
C 7	457	43.3	200707	1	BSUB0015
C 8	420.6	39.9	2664	1	BSBOPCGEN
C 9	209.6	19.9	110000	1	BA000043_25
C 10	209.6	19.9	110000	1	BA000043_26
C 11	195.4	18.5	546	6	AX433201
C 12	166	15.7	301332	1	AE017012
C 13	164.4	15.6	288814	1	AE017278
C 14	161.2	15.3	110000	1	AE017225_42
C 15	161.2	15.3	110000	1	Continuation (43 o
C 16	161.2	15.3	290029	1	AE017334_42
C 17	159.6	15.1	110000	1	AE017038
C 18	159.6	15.1	110000	1	Continuation (43 o
					Continuation (43 o

19	137	13.0	595	6	AX437557	AX437557 Sequence
C 20	132.8	12.6	290507	1	AE017327	AE017327 Listeria
21	132.2	12.5	110000	1	AP006627_16	Continuation (17 o
22	130.8	12.4	110000	1	BA000004_13	Continuation (14 o
C 23	124.8	11.8	240050	1	AL591979	AL591979 Listeria
C 24	124.8	11.8	349980	6	AX641669	AX641669 Sequence
C 25	122	11.6	239050	1	AL596169	AL596169 Listeria
C 26	122	11.6	349980	6	AX413017	AX413017 Sequence
C 27	122	11.6	349980	6	AX417045	AX417045 Sequence
28	107.6	10.2	300478	1	AE017201	AE017201 Lactobaci
29	107.6	10.2	349980	6	AX926712	AX926712 Sequence
C 30	106.8	10.1	110000	1	BA000028_20	Continuation (21 o
C 31	103.4	9.8	110000	1	CP000029_12	Continuation (13 o
C 32	103.4	9.8	300029	1	AE016748	AE016748 Staphyloc
33	98.6	9.3	110000	1	AP008934_11	Continuation (12 o
C 34	98.4	9.3	297050	1	AL935258	AL935258 Lactobaci
35	98.2	9.3	110000	1	CP000033_03	Continuation (4 of
C 36	94.2	8.9	11654	1	AE015030	AE015030 Streptoco
37	91.4	8.7	175936	14	SPNEU1908	AL449930 Streptoco
38	89.8	8.5	11344	1	AE007333	AE007333 Streptoco
39	89.8	8.5	14273	6	CQ788949	CQ788949 Sequence
40	89.8	8.5	14273	6	AR218808	AR218808 Sequence
41	89.8	8.5	14273	6	BD003720	BD003720 Polynucle
42	89.8	8.5	349980	6	AX571760	AX571760 Sequence
43	89	8.4	591	6	AX566968	AX566968 Sequence
44	89	8.4	609	6	AX617338	AX617338 Sequence
45	89	8.4	7159	6	AR354184	AR354184 Sequence

ALIGNMENTS

RESULT 1
AX930045
LOCUS AX930045 1055 bp DNA linear PAT 22-DEC-2003
DEFINITION Sequence 1 from Patent WO03087148.
ACCESSION AX930045
VERSION AX930045.1 GI:40311862
KEYWORDS
SOURCE Bacillus licheniformis
ORGANISM Bacillus licheniformis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1
AUTHORS Andersen,J.T., j Rgensen,S.T., Rasmussen,M.D., Olsen,P.B. and Clausen,I.G.
TITLE Improved bacillus host cell
JOURNAL Patent: WO 03087148-A 1 23-OCT-2003;
Novozymes A/S (DK)
FEATURES
Location/Qualifiers
source
1..1055
/organism="Bacillus licheniformis"
/mol_type="unassigned DNA"
/db_xref="taxon:1402"
34..555
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAF03678.1"
/db_xref="GI:40311863"
/translation="MYSRSKFKIGLLIGSLAALSFHLEALAKPAKVQIQLEKYYL
DGDVGINKVEAFTLDFKAGWQLIDQKGFILFRKQVDDISPLSKNTGYIGVT
EDGVISTFPHGRPGILSEPIQSFQIDIKRLSRMADDLRKGIPYRTKKEFHEHIAVK
SSGSQHVHEDMKT"

ORIGIN

Query Match 100.0%; Score 1055; DB 6; Length 1055;
Best Local Similarity 100.0%; Pred. No. 6.5e-279;
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCTTAAGCTAGGCAAAAAGCAAGTGATATTGATGACGCCGCAAGCAAGTTCAAAATC 60
|||||
Db 1 CCTTAAGTAAAGGCAAAAAGCAAGTGATATTGATGACGCCGCAAGCAAGTTCAAAATC 60
|||||

```
QY 61 GGTATTATGCTTATTGGAAGTCTGTCGCGCGCTCAGCTTTTCACCTGGAGCGCTTGCC 120
Db 61 GGTATTATGCTTATTGGAAGTCTGTCGCGCGCTCAGCTTTTCACCTGGAGCGCTTGCC 120
QY 121 GAAAGCGCGCTAAAGTTCAATCAGCTTGAAAGGTTTATCTGCGCGGAGAGCTTGGA 180
Db 121 GAAAGCGCGCTAAAGTTCAATCAGCTTGAAAGGTTTATCTGCGCGGAGAGCTTGGA 180
QY 181 ATTGAGAAATAAGTAGAGCGCGCTCGCACACTGGAAGCTTTAAAGTCTTTAAAGGG 240
Db 181 ATTGAGAAATAAGTAGAGCGCGCTCGCACACTGGAAGCTTTAAAGTCTTTAAAGGG 240
QY 241 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTTGAGCAGATT 300
Db 241 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTTGAGCAGATT 300
QY 301 TCTCCCTCAGCAAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 360
Db 301 TCTCCCTCAGCAAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 360
QY 361 TTTACGGTCCCGGGGATCTTATCAGAACCCATTCATCGTTTTCAGATTGATATA 420
Db 361 TTTACGGTCCCGGGGATCTTATCAGAACCCATTCATCGTTTTCAGATTGATATA 420
QY 421 AAGCGCTGGAAGCGCGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAA 480
Db 421 AAGCGCTGGAAGCGCGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAA 480
QY 481 AAGGAATTTGAACATGTCTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 540
Db 481 AAGGAATTTGAACATGTCTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 540
QY 541 GATATGAAGACATGACCGTGTATGCTTTTTCAGCTGCGACAGAGCTTTTTCAGCG 600
Db 541 GATATGAAGACATGACCGTGTATGCTTTTTCAGCTGCGACAGAGCTTTTTCAGCG 600
QY 601 AACATATGTTAACTTTTTCATCTAGCTTTGCTGCTTTGTTGTTACATGAAGAGCAGTC 660
Db 601 AACATATGTTAACTTTTTCATCTAGCTTTGCTGCTTTGTTGTTACATGAAGAGCAGTC 660
QY 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAAGGGACGATTGATTTATGATCGCCCC 720
Db 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAAGGGACGATTGATTTATGATCGCCCC 720
QY 721 AATATATGTTCAATTTGAAACCGCGGATCGGCTATCAGATCTTCAAGCAAAATCCGTTTA 780
Db 721 AATATATGTTCAATTTGAAACCGCGGATCGGCTATCAGATCTTCAAGCAAAATCCGTTTA 780
QY 781 TTTATAGAAAACAGCAAGAAACAACTATACATACCATTTATGTAAGAGAACACGCA 840
Db 781 TTTATAGAAAACAGCAAGAAACAACTATACATACCATTTATGTAAGAGAACACGCA 840
QY 841 ATGGCTGTACGGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATG 900
Db 841 ATGGCTGTACGGCTTTTCGACAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATG 900
QY 901 TTAAGGGATCGGCCCCAAAGAGCGCTTGCGATCTCGCTTCCGGCGATCCGGAGCGG 960
Db 901 TTAAGGGATCGGCCCCAAAGAGCGCTTGCGATCTCGCTTCCGGCGATCCGGAGCGG 960
QY 961 TGATTGAAGCGATCGAGATGAGGACGAGCAATTTCTGTCANAATTTCCCGCGCTGAGCA 1020
Db 961 TGATTGAAGCGATCGAGATGAGGACGAGCAATTTCTGTCANAATTTCCCGCGCTGAGCA 1020
QY 1021 AAAAAACGGCAAGGAGGATCATCTCTTGACCTGAAA 1055
Db 1021 AAAAAACGGCAAGGAGGATCATCTCTTGACCTGAAA 1055
```

RESULT 2
AX954657
LOCUS
DEFINITION Sequence 123 from Patent WO03093453.
AX954657 1522 bp DNA linear PAT 08-JAN-2004

```
AX954657  
AX954657.1 GI:40784029  
Bacillus licheniformis  
Bacillus licheniformis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
1  
Andersen, J.T., j Rgensen, S.T., Rasmussen, M.D., Olsen, P.B. and  
Clausen, I.G.  
Improved bacillus host cell  
Patent: WO 03093453-A 123 13-NOV-2003;  
Novozymes A/S (DK)  
FEATURES  
source  
1..1522  
/organism="Bacillus licheniformis"  
/mol_type="unassigned DNA"  
/db_xref="taxon:1402"  
501..1022  
/notes="unnamed protein product"  
/codon_start=1  
/transl_table=11  
/protein_id="CAF05523.1"  
/db_xref="GI:40784030"  
DGVGIENKVEAARTLEDKAAVKGMOLIDQKGFILFKQVDDISPLSKNGYIGVT  
EDGVISTPHCRPGILSEPIQSFPQIDIKRLESRMADDLRKGIPIYRTKKEPHEIVAVK  
SSGSQHVHEDMKT"  
ORIGIN  
Query Match 100.0%; Score 1055; DB 6; Length 1522;  
Best Local Similarity 100.0%; Pred. No. 6.7e-279;  
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTAAGGTAAGGCAAAAAGAGGTGATATTGATGTACAGCCGAGCAAGATTCAAAATC 60  
Db 468 CTTTAAGGTAAGGCAAAAAGAGGTGATATTGATGTACAGCCGAGCAAGATTCAAAATC 527  
QY 61 GGTATTATGCTTATTGGAAGTCTGTCGCGCGCTCAGCTTTTCACCTGGAGCGCTTGCC 120  
Db 528 GGTATTATGCTTATTGGAAGTCTGTCGCGCGCTCAGCTTTTCACCTGGAGCGCTTGCC 587  
QY 121 GAAAGCGCGCTAAAGTTCAATCCAGCTTGAAAGGTTTATCTGCGCGGAGAGCTTGGA 180  
Db 588 GAAAGCGCGCTAAAGTTCAATCCAGCTTGAAAGGTTTATCTGCGCGGAGAGCTTGGA 647  
QY 181 ATTGAGAAATAAGTAGAGCGCGCTCGCACACTGGAAGCTTTTAAAGTCTTTTAAAGGG 240  
Db 648 ATTGAGAAATAAGTAGAGCGCGCTCGCACACTGGAAGCTTTTAAAGTCTTTTAAAGGG 707  
QY 241 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTTGAGCAGATT 300  
Db 708 TGGCAGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTTGAGCAGATT 767  
QY 301 TCTCCCTCAGCAAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 360  
Db 768 TCTCCCTCAGCAAAAACAAACGGTTTATATCGGAGTGACTGAAGATGGCGTGAATTCGACT 827  
QY 361 TTTACCGGTCCCGGGGATCTTATCAGAACCCATTCATCGTTTTCAGATTGATATA 420  
Db 828 TTTACCGGTCCCGGGGATCTTATCAGAACCCATTCATCGTTTTCAGATTGATATA 887  
QY 421 AAGCGCTGGAAGCGCGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAAA 480  
Db 888 AAGCGCTGGAAGCGCGATGCGGATGATCTGCGCAAGGATACCATACCGCACGAAA 947  
QY 481 AAGGAATTTGAACATGTCTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 540  
Db 948 AAGGAATTTGAACATGTCTTGAAGCCGTAATAATCATCCGGAAGCAACATCATGTAGAA 1007  
QY 541 GATATGAAGACATGACCGTGTATGCTTTTTCAGCTGCGACAGAGCTTTTTCAGCG 600  
Db 1008 GATATGAAGACATGACCGTGTATGCTTTTTCAGCTGCGACAGAGCTTTTTCAGCG 1067
```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

CDS

ORIGIN

QY 601 AACATATGTTAACTTTTTCATCTAGCTTTCGCTGTTTGTGTTTCAATGAAGAGCAGTC 660
Db 1068 AACATATGTTAACTTTTTCATCTAGCTTTCGCTGTTTGTGTTTCAATGAAGAGCAGTC 1127
QY 661 AAAGAGGTGAATGAACGTTTGAATTCGTAAGAGGACGATTGATTATGTTATCGCCCC 720
Db 1128 AAAGAGGTGAATGAACGTTTGAATTCGTAAGAGGACGATTGATTATGTTATCGCCCC 1187
QY 721 AATATATTTGTCATTTGAAACCGCGGATCGGCTATFCAGATCTTACGCCAAATCCGTTT 780
Db 1188 AATATATTTGTCATTTGAAACCGCGGATCGGCTATFCAGATCTTACGCCAAATCCGTTT 1247
QY 781 TTTTATAAGAAAACAGCAAGAAACAATCTATACATACATATGTTAGAGAGACAGCA 840
Db 1248 TTTTATAAGAAAACAGCAAGAAACAATCTATACATACATATGTTAGAGAGACAGCA 1307
QY 841 ATGCGCTGTCAGGCTTTTCGCAAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATG 900
Db 1308 ATGCGCTGTCAGGCTTTTCGCAAGGGAAGAAAATGCTGTTTACGAAAATGCTGAATG 1367
QY 901 TTACGGGATCGGCCCAAAAGAGCGCTTTCGATCTTCGCTTCCGGCGATCCGGAGCGG 960
Db 1368 TTACGGGATCGGCCCAAAAGAGCGCTTTCGATCTTCGCTTCCGGCGATCCGGAGCGG 1427
QY 961 TGATTGAAGCGATCGAGAAATGAGGACGAAAGCAATTTCTCGTCAAAATTTCCCGCGGTAGGCA 1020
Db 1428 TGATTGAAGCGATCGAGAAATGAGGACGAAAGCAATTTCTCGTCAAAATTTCCCGCGGTAGGCA 1487
QY 1021 AAAAAGCGCAAGGAGATCATCTTTGACCTGAAA 1055
Db 1488 AAAAAGCGCAAGGAGATCATCTTTGACCTGAAA 1522

RESULT 3

AE017333_27/c

WPCOMMENT

Sequence split into 43 fragments LOCUS AE017333 Accession AE017333

Fragment Name	Begin	End
AE017333_00	1	110000
AE017333_01	100001	210000
AE017333_02	200001	310000
AE017333_03	300001	410000
AE017333_04	400001	510000
AE017333_05	500001	610000
AE017333_06	600001	710000
AE017333_07	700001	810000
AE017333_08	800001	910000
AE017333_09	900001	1010000
AE017333_10	1000001	1110000
AE017333_11	1100001	1210000
AE017333_12	1200001	1310000
AE017333_13	1300001	1410000
AE017333_14	1400001	1510000
AE017333_15	1500001	1610000
AE017333_16	1600001	1710000
AE017333_17	1700001	1810000
AE017333_18	1800001	1910000
AE017333_19	1900001	2010000
AE017333_20	2000001	2110000
AE017333_21	2100001	2210000
AE017333_22	2200001	2310000
AE017333_23	2300001	2410000
AE017333_24	2400001	2510000
AE017333_25	2500001	2610000
AE017333_26	2600001	2710000
AE017333_27	2700001	2810000
AE017333_28	2800001	2910000
AE017333_29	2900001	3010000
AE017333_30	3000001	3110000
AE017333_31	3100001	3210000
AE017333_32	3200001	3310000
AE017333_33	3300001	3410000
AE017333_34	3400001	3510000

AE017333_35 3500001 3610000
AE017333_36 3600001 3710000
AE017333_37 3700001 3810000
AE017333_38 3800001 3910000
AE017333_39 3900001 4010000
AE017333_40 4000001 4110000
AE017333_41 4100001 4210000
AE017333_42 4200001 4222645
Continuation (28 of 43) of AE017333 from base 2700001 (AE017333 Bacillus licheniformis D

Query Match 99.3%; Score 1047.4; DB 1; Length 110000;
Best Local Similarity 99.9%; Pred. No. 1.le-276;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 GGTAAAGCAAAAAGAGGTGATATTGATGTACAGCCGAAAGCAAGTTCAAAATCCGTTT 66
Db 92432 GTTAAAGCAAAAAGAGGTGATATTGATGTACAGCCGAAAGCAAGTTCAAAATCCGTTT 92373
QY 67 TTGCTTATTGAAAGTCTGCTGGCGCGCTCAGCTTTTCACTGGAGGCCTTTGGCCGAAAAG 126
Db 92372 TTGCTTATTGAAAGTCTGCTGGCGCGCTCAGCTTTTCACTGGAGGCCTTTGGCCGAAAAG 92313
QY 127 CCGGCTAAAGTTCAAATCCAGCTTTGAAAGGTTTATCTGGACGGAGAGCTTGGAAATTGAG 186
Db 92312 CCGGCTAAAGTTCAAATCCAGCTTTGAAAGGTTTATCTGGACGGAGAGCTTGGAAATTGAG 92253
QY 187 AATAAGTAGAGGCGCTCGCACACTGGAAGACTTTTAAAGCTGCTTATAAAGGGTGGCAG 246
Db 92252 AATAAGTAGAGGCGCTCGCACACTGGAAGACTTTTAAAGCTGCTTATAAAGGGTGGCAG 92193
QY 247 CTCAATCGATCAGAAAAAGGGGTTTATTCTGTTTCGCAAAACAGGTGGACGACATTTCTCCC 306
Db 92192 CTCAATCGATCAGAAAAAGGGGTTTATTCTGTTTCGCAAAACAGGTGGACGACATTTCTCCC 92133
QY 307 CTCAAGCAAAACAAACGGTTTATCGAGTACTGAGAGTGGCGTGATTTGCTTTTTCAC 366
Db 92132 CTCAGCAAAACAAACGGTTTATCGAGTACTGAGAGTGGCGTGATTTGCTTTTTCAC 92073
QY 367 GGTCCCGCGGCATCTTATCAGAAACCCATTCATTCGTTTTCAGATTGATATAAGCGG 426
Db 92072 GGTCCCGCGGCATCTTATCAGAAACCCATTCATTCGTTTTCAGATTGATATAAGCGG 92013
QY 427 CTGAAAACCGCGATGGCGGATGATCTGCGCAAAAGGATACCATCCGACCGAAAGGAA 486
Db 92012 CTGAAAACCGCGATGGCGGATGATCTGCGCAAAAGGATACCATCCGACCGAAAGGAA 91953
QY 487 TTTGAAATGTCATTGAAGCCGTAAATCATCCGGAAGCCAAACATCATGTAGAGATATG 546
Db 91952 TTTGAAATGTCATTGAAGCCGTAAATCATCCGGAAGCCAAACATCATGTAGAGATATG 91893
QY 547 AAGACATGACGCTGTTATGCTTTTTCAGCTGCGAGACAGAGCTTTTTCAGCGAACATA 606
Db 91892 AAGACATGACGCTGTTATGCTTTTTCAGCTGCGAGACAGAGCTTTTTCAGCGAACATA 91833
QY 607 TGTTAACTTTTTCATTTAGCTTTGCTTTCGTTTGTGTTTACAAATGAAGAGCAGTCAAAGAG 666
Db 91832 TGTTAACTTTTTCATTTAGCTTTGCTTTCGTTTGTGTTTACAAATGAAGAGCAGTCAAAGAG 91773
QY 667 GTGAATGAACGTTGATCGAATTCGTAAGGGACGATGATGATGATCGCCCCCAATATA 726
Db 91772 GTGAATGAACGTTGATCGAATTCGTAAGGGACGATGATGATGATGATCGCCCCCAATATA 91713
QY 727 TTGTCAATTTGAAAACGGCGGATCGGCTATCAGATCTTTCAGCCCAAAATCCGTTTATTATA 786
Db 91712 TTGTCAATTTGAAAACGGCGGATCGGCTATCAGATCTTTCAGCCCAAAATCCGTTTATTATA 91653
QY 787 AGAAAAACAGCAAAAGAAACCAATCTATACATACATACATGTTAGAGAGACGAGTGGC 846
Db 91652 AGAAAAACAGCAAAAGAAACCAATCTATACATACATACATGTTAGAGAGACGAGTGGC 91593
QY 847 TGTACGGCTTTTCAGCAAGGGAGAAAAAATGCTGTTTACGAAAATGCTGAATGTTACGG 906
Db 91592 TGTACGGCTTTTCAGCAAGGGAGAAAAAATGCTGTTTACGAAAATGCTGAATGTTACGG 91533

Qy	907	GGATCGGCCCAAGAGCGCTTCGGATCCTCGCTTCGGCGATCCGGAGCGGTGATTG	966
Db	91532	GGATCGGCCCAAGAGCGCTTCGGATCCTCGCTTCGGCGATCCGGAGCGGTGATTG	91473
Qy	967	AAGCGATCGAAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGGGCGTAGGCCAAAAAA	1026
Db	91472	AAGCGATCGAAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGGGCGTAGGCCAAAAAA	91413
Qy	1027	CGGCAAGCGAGATCATCTTGAACCTGAAA	1055
Db	91412	CGGCAAGCGAGATCATCTTGAACCTGAAA	91384
RESULT 4			
CP000002_27/c			
WPCOMMENT			
Sequence split into 43 fragments LOCUS CP000002 Accession CP000002			
Fragment Name Begin End			
CP000002_00 1 110000			
CP000002_01 100001 210000			
CP000002_02 200001 310000			
CP000002_03 300001 410000			
CP000002_04 400001 510000			
CP000002_05 500001 610000			
CP000002_06 600001 710000			
CP000002_07 700001 810000			
CP000002_08 800001 910000			
CP000002_09 900001 1010000			
CP000002_10 1000001 1110000			
CP000002_11 1100001 1210000			
CP000002_12 1200001 1310000			
CP000002_13 1300001 1410000			
CP000002_14 1400001 1510000			
CP000002_15 1500001 1610000			
CP000002_16 1600001 1710000			
CP000002_17 1700001 1810000			
CP000002_18 1800001 1910000			
CP000002_19 1900001 2010000			
CP000002_20 2000001 2110000			
CP000002_21 2100001 2210000			
CP000002_22 2200001 2310000			
CP000002_23 2300001 2410000			
CP000002_24 2400001 2510000			
CP000002_25 2500001 2610000			
CP000002_26 2600001 2710000			
CP000002_27 2700001 2810000			
CP000002_28 2800001 2910000			
CP000002_29 2900001 3010000			
CP000002_30 3000001 3110000			
CP000002_31 3100001 3210000			
CP000002_32 3200001 3310000			
CP000002_33 3300001 3410000			
CP000002_34 3400001 3510000			
CP000002_35 3500001 3610000			
CP000002_36 3600001 3710000			
CP000002_37 3700001 3810000			
CP000002_38 3800001 3910000			
CP000002_39 3900001 4010000			
CP000002_40 4000001 4110000			
CP000002_41 4100001 4210000			
CP000002_42 4200001 422334			
Continuation (28 of 43) of CP000002 from base 2700001 (CP000002 Bacillus licheniformis A			
Query Match 99.3%; Score 1047.4; DB 1; Length 110000;			
Best Local Similarity 99.9%; Pred. No. 1.1e-276;			
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	7	GGTAAGCCAAAAGAGGATGATATTGATGTACAGCCGAGCAAGTTCAAATCGGTTA	66
Db	91570	GTTAAGGCCAAAAGAGGATGATATTGATGTACAGCCGAGCAAGTTCAAATCGGTTA	91511
Qy	67	TTGCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTACCTGGAGGCGCTTGGCCGAAAAAG	126

RESULT 5
AX434675
LOCUS
DEFINITION Sequence 3090 from Patent WO0229113.
ACCESSION AX434675

AX434675
459 bp
DNA
linear
PAT 28-JUN-2002

VERSION	AX434675.1	GI:21659483	
KEYWORDS	Bacillus licheniformis		
SOURCE	Bacillus licheniformis		
ORGANISM	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
REFERENCE	1	Berka, R. and Clausen, I. G.	
AUTHORS	Methods for monitoring multiple gene expression		
TITLE	Patent: WO 0229113-A 3090 11-APR-2002;		
JOURNAL	Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)		
FEATURES	Location/Qualifiers		
source	1. 459		
	/organism="Bacillus licheniformis"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:1402"		
ORIGIN			
Query Match	43.5%;	Score 459;	DB 6; Length 459;
Best Local Similarity	100.0%;	Pred. No. 7.1e-115;	
Matches 459;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	65	TATTCTTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTTCACTGGAGCGCTTGGCCGAAA	124
Db	1	TATTGCTTATTGGAAGTCTGCTGGCGCGCTCAGCTTTTCACTGGAGCGCTTGGCCGAAA	60
Qy	125	AGCGCGCTAAAGTTCAAATCCAGCTTGAAGAGTTTATCTGACGCGACGCTTGGAAATTG	184
Db	61	AGCGCGCTAAAGTTCAAATCCAGCTTGAAGAGTTTATCTGACGCGACGCTTGGAAATTG	120
Qy	185	AGATAAAGTAGAGCGCGCTCGCACACTGGAAGCTTTAAAGCTGCTTATAAAGGTTGC	244
Db	121	AGATAAAGTAGAGCGCGCTCGCACACTGGAAGCTTTAAAGCTGCTTATAAAGGTTGC	180
Qy	245	AGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAACAGGTGACGACATTTCCTC	304
Db	181	AGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAACAGGTGACGACATTTCCTC	240
Qy	305	CCCTCAGCAAAACAAACCGTTATATCGAGTGACTGAAGATGGCGTGAATTCGACTTTTC	364
Db	241	CCCTCAGCAAAACAAACCGTTATATCGAGTGACTGAAGATGGCGTGAATTCGACTTTTC	300
Qy	365	ACGGTCGCGCGGCACTTATCAGAACCCATCAATCGTTTTTCAGATTGATATAAGC	424
Db	301	ACGGTCGCGCGGCACTTATCAGAACCCATCAATCGTTTTTCAGATTGATATAAGC	360
Qy	425	GCCTGGAAGCGGATGCGGATGATCTGCGCAAGAGGATACCATACCGCAGCAAAAAGG	484
Db	361	GCCTGGAAGCGGATGCGGATGATCTGCGCAAGAGGATACCATACCGCAGCAAAAAGG	420
Qy	485	AATTTGAACATGCTATTGAAGCCGTAAATCATCCGGAA	523
Db	421	AATTTGAACATGCTATTGAAGCCGTAAATCATCCGGAA	459
RESULT 6			
BSY15896			
LOCUS			
DEFINITION	Bacillus subtilis nado, yrbA, yrbB, yrbC, yrbD, orf7, yrbE, csbX, bofC, ruvA, ruvB, queA, tgt, yrbF, orf16, yrbG, spoVB genes and partial nadC gene.		
ACCESSION	Y15896		
VERSION	Y15896.1		
KEYWORDS	bofC gene; csbX gene; nado gene; nadC gene; nicotinate-nucleotide pyrophosphorylase; queA gene; quinolinate synthetase; ruvA gene; ruvB gene; sporulation-specific protein; spoVB gene; tgt gene; tRNA-guanine transglycosylase; yrbA gene; yrbB gene; yrbC gene; yrbD gene; yrbE gene; yrbF gene; yrbG gene.		
SOURCE	Bacillus subtilis		
ORGANISM	Bacillus subtilis		
REFERENCE	1	Tosato, V., Bolotin, A., Bertani, I., Valentino, I., and Bruschi, C.V.	
AUTHORS	A 17.8 kb segment in the spoVB-nadC region of the Bacillus subtilis		
TITLE			

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 17787)
AUTHORS	Bruschi, C.V.
TITLE	Direct Submission
JOURNAL	Submitted (11-DEC-1997) C.V. Bruschi, International Centre for Genetic Engineering and Biotechnology, Padriciano 99, I-34012 Trieste, ITALY
FEATURES	Location/Qualifiers
source	1. 17787
	/organism="Bacillus subtilis"
	/mol_type="genomic DNA"
	/strain="168"
	/db_xref="taxon:1423"
	/map="242-243"
	<1. .167
	/gene="nadC"
	<1. .167
	/gene="nadC"
	/function="NAD biosynthesis"
	/codon_start=3
	/transl_table=11
	/product="nicotinate-nucleotide pyrophosphorylase"
	/protein_id="CAB75320.1"
	/db_xref="GI:6977795"
	/db_xref="GOA:Q9KWZ2"
	/db_xref="InterPro:IPR002638"
	/db_xref="UniProt/TREMBL:Q9KWZ2"
	/translation="ASGGITLESIPAFKGTGVNYSIGLFLTHSVKSLDISMDVTLSNE
	SVECCTVNS"
	169. .1260
	/gene="nada"
	169. .1260
	/gene="nada"
	/function="NAD biosynthesis"
	/codon_start=1
	/transl_table=11
	/product="quinolinate synthetase A"
	/protein_id="CAB75321.1"
	/db_xref="GI:6977796"
	/db_xref="GOA:Q9KWZ1"
	/db_xref="UniProt/Swiss-Prot:Q9KWZ1"
	/translation="MIKQSDNMPSEYKELSRKDMETRVAAIKKFGSRLFIPIGHYQ
	KDEVIQADQDGTDSLQAVAKNKEDYIVFCGVHPMAETADMLTSSQQTWVLPDMR
	AGSMADMAQOTNRANKLQHFIDTIIPLTYVNSTABIKAFVKGHGATVTSNA
	KVLWAPTKKRILFLPDQHLGRNTAYDLGIALEDMAVDPKMDLVAESGHTNVKV
	ILWKGHCSVHEKTTKNIHDMREDPDQIIIVHPECSEVVTLSDDNGSTKYIIDTN
	QAPGSGWAGTGMNIVQRIIHEHPDKQIESLNPDMPCLTMNRIDILPHLLWSLEQIE
	KGPESGVKVPKAIQEDALLANRMLSI"
	1377. .2540
	/gene="yrbA"
	1377. .2540
	/gene="yrbA"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="CAB75322.1"
	/db_xref="GI:6977797"
	/db_xref="GOA:O32062"
	/db_xref="InterPro:IPR002482"
	/db_xref="UniProt/TREMBL:O32062"
	/translation="MKIHIQKGSLSMKIAEKYGVDEVVKLNTQLSNPDLIMPGMK
	IKVPSGVPRKEPKAGSPAGSVKQEHVPAKEKPSVVDVEDTKPKKESMPYVP
	MPNLQENVYADVNDYDMKQLFQWSPKPKPKKHGDMHDMHYHMQDFPQOE
	MSNMENYNNMNPMPKAPVPGIIEENVHTVNMMPAPQPYHYHFAHFPVFPV
	SPILPGLGCLYPPYPAQYPMHPMHGYPQVSPQYDPGYENQHSHHGHGYSYGA
	PQYASPAYSPYGHMPYGPYPTQVNGVQYPAAGYMPYKDHDDCCDGDHQPYS
	APGSHGMGAYGSPNNPYGTANPNPNPYSGVSMPTNQPSVNMFGRESENE"
	2766. .3284
	/gene="yrbB"
	2766. .3284
	/gene="yrbB"
	/codon_start=1

```
gene
CDS
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75323.1"
/db_xref="GI:6977798"
/db_xref="GOA:P94446"
/db_xref="UniProt/Swiss-Prot:P94446"
/transl_table="MNTTRNNGNTRPIGYTNNENADRQGDGIDHDGPFVSELMDQND
GNRTTNNNRDRVTADDRVPLATDGTNNNNMNRNNAANNQYDQENRRLAAKTA
NRUKQKNVNDQVMVSDRRVVIKSHREFTKSDRNVVKARNYANGRDVQVSTDK
GLRKLHMMNR"
3387..4109
/gene="yrbC"
3387..4109
/gene="yrbC"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75324.1"
/db_xref="GI:6977799"
/db_xref="GOA:P94447"
/db_xref="InterPro:IPR002876"
/db_xref="UniProt/Swiss-Prot:P94447"
/transl_table="MAGHSKWKNIQKKNQAQDARKGI FMKLAKIYVAAKEGGADPE
ANSALRVIEKAGAMPNENIDRAIKKAGGQDGSSEYETTYEGYGPSTYAMVECV
TDNKNRTASNVRTPAFNKGSGSLGESGCVAFIFERKGFITIDRTERQIEEDLMDVLE
AGGELRIEEDLYXEVTEPEHFEVTKALBSKYLISAEVTLNPTVTAEDDQAVEKL
ETLLDVEDDDDDQVEVTVNRS"
complement(4150..5604)
/gene="yrbD"
complement(4150..5604)
/gene="yrbD"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75325.1"
/db_xref="GI:6977800"
/db_xref="GOA:O32060"
/db_xref="InterPro:IPR001463"
/db_xref="InterPro:IPR002293"
/db_xref="UniProt/TrEMBL:O32060"
/transl_table="MADFVSLNVLWSTPVYIYLLGIGFAPFSIMTRFLQVRHLKEMI
VQMFKSSSEAGVSFALSIALSRVGTGNIAGVATAIAFGPGAVFWMALFPIGA
ASAVESTLAQYVKVQGYRGPRVYIEKGLIKWFAVLFAAAALIAMAFILMPGVQ
SNSAAGIONAFGISPVFTGCGVLVLLGFIIFGVKRIANNAOMIVPMAIGYLLSL
IILVNVSELPAVLSLIFKSAFALDSAFGLIGMAISGWKRGYISNEAGCGTGPHPA
AAAEVSRPKQGLVQARSFYIDTLFVCSATAFMILFTGMVNTQAADGSFIVHQLKVE
AGPGFTQAADSVLPFGGAGEVAIALFFAFTTIMYYIYAETNIAYIARGRESKWAM
LGLKLIILAAATFYGTAKTASLAWALGDAGLGMVNLVIALVILLAKPALLALKDYERQ
KKOGLDPIFDPKALGIKNADFEKEYTHESRVS"
complement(6712..7104)
/note="orf7"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75326.1"
/db_xref="GI:6977801"
/db_xref="GOA:O32057"
/db_xref="InterPro:IPR000719"
/db_xref="UniProt/TrEMBL:O32057"
/transl_table="MLTKAEALAHTVYRKNSRFPDKVKRKESEELTLICKGRSAYVFAL
TEGGRKVALKVFPEYQATAVKEAAIYEKLAGSASYPDIYETGDSFILMEYIKETFY
NCLKKGIATISDMIQQVEEALSDAQPD"
7342..8367
/gene="yrbE"
7342..8367
/gene="yrbE"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75327.1"
/db_xref="GI:6977802"
/db_xref="GOA:O05389"
/db_xref="InterPro:IPR000683"
gene
CDS
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB75328.1"
/db_xref="GI:6977803"
/db_xref="GOA:O05390"
/db_xref="InterPro:IPR004748"
/db_xref="InterPro:IPR007114"
/db_xref="UniProt/Swiss-Prot:O05390"
/transl_table="MNTVHAKGNLKNKIGIPSHMWGYIGVVIEMVGDGLBQGLWSPF
LVDRGLSMQSGASLFTMYGTAVTISAWLSGTFTVQTPRKMTVGLLAFIIGSAAFIG
WAPHMVYPALLGSLYALRGYPLFYSLVFWVSYSYSONILGAVGFWFMFTCGLN
VLGPFYSSYAVPAFGEINTLSALLFVAAGSILALFNKOKFTPIQKODQPKWELSK
AFTIMENPKVGGVGVKTNIALGQGFALFPLTYLARYGYSVSENLQINGTLFVNI
VFNIIFGAVDGLGNRWNTVMVFGVGGGIFTLALITYTPQLIGHQYVWLMIIACCYGAA
LAGYVPLSALLPTLAPDNKGAASVNLGFLCAPIAGIVSLFGLPGAGGVWIWFA
ALYFFSAFLTRFLTISQSTDVYTEERFVRENVQTNFDKTVKQ"
9793..10305
/gene="bofC"
9793..10305
/gene="bofC"
/codon_start=1
/transl_table=11
/product="BofC protein"
Query Match 43.3%; Score 457; DB 1; Length 17787;
Best Local Similarity 69.0%; Pred. No. 3.2e-114;
Matches 659; Conservative 0; Mismatches 285; Indels 11; Gaps 2;
Qy 101 TTCACTCTGGAGCGCTTGGCCGMAAAGCGGCTAAAGTTCAAATCCAGCTTCAAAAGGTTT 160
Db 9878 TTGGCGCGGAAGTGGACACATATGAGCGCGTGCAGTGCATGTCGAGCTTGAAGAAGTGT 9937
Qy 161 ATCTGGAAGGAGCGTGGGAATTGAGAAATAGAGAGCGCGCTCGCACACTCGAAGACT 220
Db 9938 ATTTGGACGGCGATGTGAGCATTCAGCATTAAGCATGAAAGAGTGTCTTCTATGATGATT 9997
Qy 221 TTAAGCTGCTTATAAGGGTGGCAGCTCATCGATCAGAAAAGGGGTTTATTCTGTTTC 280
Db 9998 TTTGGCGAGCTTATGCGCGATGGACGCTAGTGGAAACAAAAGGGATATGTCTCTTCC 10057
Qy 281 GCAACAGGTGGACGACATTTCTCCCTCAGCAAAAACAAACGGTTATATCGAGTGACTG 340
Db 10058 GGAAGCAAAATGGACGATATTTCTCCGCTCAGCAAGGTCAACGGGTATATCGGTATCGG 10117
Qy 341 AAGATGGCGTGATTTGACATTTTTCACGGTCCCGCGGCGCATCTTATCAGAACCCATTCAAT 400
Db 10118 ATATGGAGTGATTTCCACTTTTTCATGGCGGCGCTGAGCCAGCTTCCGAACCAATTCAGT 10177
Qy 401 CGTTTTTTCAGATTGATATAAGCGGCTGGAAAGCGGATGGCGGATGATCTCGCAAG 460
Db 10178 CTTTTTTTCAAATTGATTTAGAAAGGCTGGAAAGCCCATATGCAAAAAAATCTGCTGAAAG 10237
Qy 461 GGATACCATACCGCACGAAAGGAATTGAAATGTCATTGAAGCCGTAAATATCATCCG 520
Db 10238 GCATTCATTTCCGACAAAGCGAGTTTGAGGATGTCATCGAA-----CATATG 10287
Qy 521 GAAGCCCAACATCATGTATAGAGATATGAAGACATGACCGCTGTATGTCTTTTTCAGCTGC 580
Db 10288 AAGACATACAGCGGTAGCGTCCGCTGATTGAAGACATGATTATTCATGCTCTTTTTCGT 10347
```

```
gene
CDS
/db_xref="InterPro:IPR004104"
/db_xref="UniProt/Swiss-Prot:O05389"
/transl_table="MTYQIVTGIIGAGRIKGLHVQNIISRPHMKIKAISDIQASRIKS
WADSHQIEYITSDIRLLDDPIDAIFCPTAVHQAOMIKAEAAKHHIPCEKPVSPS
LDETSEALAAVRKHGVTLOQFNRRFDPHFKKIKTIVEENGIEIGTPHLLKITSRDPEPP
NIDYVRTSGFLPMDMSIHDEDMARYIMGSEVTEYAKGAALVNSPFAELGDIDTAVIT
LTPENGAMVIDNSROAVGYDQRVFVGTKGSAADNSRPTTVEVSTADPVMKDKPH
FFFLERYKDSYEBEILLFAEAI GTNQETPCTGNDGLQAGRIARAQQSLAFGMPVSI
HTSKIAP"
8438..9745
/gene="cebX"
8438..9745
/gene="cebX"
/codon_start=1
/transl_table=11
/product="cebX protein"
/protein_id="CAB75328.1"
/db_xref="GI:6977803"
/db_xref="GOA:O05390"
/db_xref="InterPro:IPR004748"
/db_xref="InterPro:IPR007114"
/db_xref="UniProt/Swiss-Prot:O05390"
/transl_table="MNTVHAKGNLKNKIGIPSHMWGYIGVVIEMVGDGLBQGLWSPF
LVDRGLSMQSGASLFTMYGTAVTISAWLSGTFTVQTPRKMTVGLLAFIIGSAAFIG
WAPHMVYPALLGSLYALRGYPLFYSLVFWVSYSYSONILGAVGFWFMFTCGLN
VLGPFYSSYAVPAFGEINTLSALLFVAAGSILALFNKOKFTPIQKODQPKWELSK
AFTIMENPKVGGVGVKTNIALGQGFALFPLTYLARYGYSVSENLQINGTLFVNI
VFNIIFGAVDGLGNRWNTVMVFGVGGGIFTLALITYTPQLIGHQYVWLMIIACCYGAA
LAGYVPLSALLPTLAPDNKGAASVNLGFLCAPIAGIVSLFGLPGAGGVWIWFA
ALYFFSAFLTRFLTISQSTDVYTEERFVRENVQTNFDKTVKQ"
9793..10305
/gene="bofC"
9793..10305
/gene="bofC"
/codon_start=1
/transl_table=11
/product="BofC protein"
```


Qy	641	TGTTTCAATGAAGACAGTCAAGAGGTGAATGAACCTTGATCGAATTCGTAAAGGGAC	700
Db	2313	TGGTACAAATGAAGAACAGTCAAGAGGTGAATTTGCG-TGATTCGAATTTGTTAAAGGGAC	2371
Qy	701	GATTGATTATGTATCGCCCCCAATATATTGTCATTGAAACCGCGGGATCGGTATCAGAT	760
Db	2372	GATTGATTATGTATCTCCCCAATATATTGTATAGAACCGCGGCATCGGTATCAGAT	2431
Qy	761	CTTCACGCCAAATCCGTTTATTTATAAGAAAAACAGCAAGAAACAATCTATACATACCA	820
Db	2432	TTTCACGCCGAATCCGTTTATATATAAGGAGAGAACCAAGAAACGATTTTACATACCA	2491
Qy	821	TTATGTTAAGAGAGACACAGAAATGGCGTGTACGGCTTTTTCGACAGGGGAAGAAAAATGCT	880
Db	2492	TCATATCAGAGAGGATGCTTTTTCGCTTTTATGTGTTTTTCAACAGAGAGAGAGAGGCGCT	2551
Qy	881	GTTTACGAAAATGCTGTAATGTACGGGGATCGGCCCAAAAGGAGCGCTTCGCAATCCTCGC	940
Db	2552	CTTTTACAAAGCTGTTAAACGTGACCGGAATTGGCCCGAAGGAGCATTTGCCATCTTTGGG	2611
Qy	941	TTCCGGCGATCCGGGACGGGTGATTGAAGCGATCGAGAATGAGGACGAAGCAT	993
Db	2612	CTCGGTTGATCCCGGAGCGGTATCCAAAGCAATTGAAAAATGAAGATCGAAGCTT	2664

RESULT 9

RESULI 9 25/0
BA000043 25/0

BA000043
WPCOMMENT

WPCOMMENT

Fragment Name	Begin	End
BA000043_00	1	110000
BA000043_01	100001	210000
BA000043_02	200001	310000
BA000043_03	300001	410000
BA000043_04	400001	510000
BA000043_05	500001	610000
BA000043_06	600001	710000
BA000043_07	700001	810000
BA000043_08	800001	910000
BA000043_09	900001	1010000
BA000043_10	1000001	1110000
BA000043_11	1100001	1210000
BA000043_12	1200001	1310000
BA000043_13	1300001	1410000
BA000043_14	1400001	1510000
BA000043_15	1500001	1610000
BA000043_16	1600001	1710000
BA000043_17	1700001	1810000
BA000043_18	1800001	1910000
BA000043_19	1900001	2010000
BA000043_20	2000001	2110000
BA000043_21	2100001	2210000
BA000043_22	2200001	2310000
BA000043_23	2300001	2410000
BA000043_24	2400001	2510000
BA000043_25	2500001	2610000
BA000043_26	2600001	2710000
BA000043_27	2700001	2810000
BA000043_28	2800001	2910000
BA000043_29	2900001	3010000
BA000043_30	3000001	3110000
BA000043_31	3100001	3210000
BA000043_32	3200001	3310000
BA000043_33	3300001	3410000
BA000043_34	3400001	3510000
BA000043_35	3500001	3544776

01 JULISA

RESULT TO
B3000043 36/5

BA000043
WPCOMMENT

WPCOMMENT

Fragment Name	Begin	End
BA000043_00	100001	110000
BA000043_01	100001	210000
BA000043_02	300001	310000
BA000043_03	300001	410000
BA000043_04	400001	510000
BA000043_05	500001	610000
BA000043_06	600001	710000
BA000043_07	700001	810000
BA000043_08	800001	910000
BA000043_09	900001	1010000
BA000043_10	1000001	1110000
BA000043_11	1100001	1210000
BA000043_12	1200001	1310000
BA000043_13	1300001	1410000
BA000043_14	1400001	1510000
BA000043_15	1500001	1610000
BA000043_16	1600001	1710000
BA000043_17	1700001	1810000
BA000043_18	1800001	1910000
BA000043_19	1900001	2010000
BA000043_20	2000001	2110000
BA000043_21	2100001	2210000
BA000043_22	2200001	2310000
BA000043_23	2300001	2410000
BA000043_24	2400001	2510000
BA000043_25	2500001	2610000
BA000043_26	2600001	2710000
BA000043_27	2700001	2810000
BA000043_28	2800001	2910000
BA000043_29	2900001	3010000
BA000043_30	3000001	3110000
BA000043_31	3100001	3210000
BA000043_32	3200001	3310000
BA000043_33	3300001	3410000
BA000043_34	3400001	3510000
BA000043_35	3500001	3547776

Continuation (27 of 36) of BA0000043 from base 2600001 (BA0000043 Geobacillus kaustophilus

Query Match	19.9%;	Score 209.6;	DB 1;	Length 110000;
Best Local Similarity	54.0%;	Pred. No. 4.6e-46;		
Matches	528;	Conservative	0;	Mismatches 404; Indels 46; Gaps 3;
Qy	84	GCTGCCCGCTCAGCTTTTCACTGGAGGCTTGGCGGAAAGCGGCTTCAAGTTCAAT	143	
Db	8986	GATGCCGCCATCGTGTTCGGTTCATTCGGCTTCGCGCACCGGTTAAGATGACGAT	8927	
Qy	144	CCAGCTTGAAGAAGTTTATCTGGACGAGAGCTTGGAAATGAGATAAAGTAGAGCGCG	203	
Db	8926	TGTGCTGAGCGGCAATACTTTGGACGAGAGATGAGCGAAGAAAGTCAACCGAAACCGT	8867	
Qy	204	TGCGACACTGGAAGACTTTAAAGCTGCTTATAAAGGTGGAGCTTCATCGATCAGAAAA	263	
Db	8866	CGATTGCGATGACAGAAATATGGAATAATTCGGGTGGCACTCGTACGCTCGATGA	8807	
Qy	264	GGGGTTTATCTGTTTCGCAACAGGTTGGACGACATTTCTCCCTCAGCAAAAACAAACGG	323	
Db	8806	TCAGACGATCGTATTTTCGCAAAACAAATCAACGACATTTTCGGCGTTGTTTAAACAAACGG	8747	
Qy	324	TTATATCGGAGTACTGAAGATGGGTGATTTTCGACTTTTCACGGTCCCGCGGATCTT	383	
Db	8746	CTACTTCGGAATACGGACGATGGCACGTCTGTCTATCTTCAACGCGCAACCGCGCGTC	8687	
Qy	384	ATCAGAACCCATTCATCTGTTTTCAGATTGATATAAGCGGTGGAAAGCCGATGGC	443	
Db	8686	GAGCGAATCATCTGTTTTCCTGTTTTCCTGTTTTCCTGTTTTCCTGTTTTCCTGTTT	8627	
Qy	444	GGATGATCTGGCAAGGATACCATACCGCACGAAAGAAATTTGAACATGTCATTGA	503	
Db	8626	AGAAAGCTGAAGAAAGGAATTCGGCTCTCTCGAAGAGCGGTATGAACAGTATCGA	8567	
Qy	504	AGCCGTAATATCTCCGGAAGCAACATCATGTGATGAGATATGAGACATGACGCTGTTA	563	

Db	8566	A-----ATGTACCGCCATTTTTCGGTCG	8544	
Qy	564	TGTCTTTTTCAGCTGCAGACAGAAAGCTTTTTCAGCAACATATGTTAACTTTTTCATTC	623	
Db	8543	T---CCATATAACCAAGCGAAAGCCGTGGTGCAGAGCGCATCGCGGATTTCTTTTTCG	8487	
Qy	624	TAGCTTTGCTGTTTGTGTTTACATCAATGAGAGCAGTCAAGAGAGTGAATGAACGTTGATC	683	
Db	8486	TATAATGAATCATATGATACATCAATGAAGAGCAGAAAGAGGAGAGGATTCATTTGATC	8427	
Qy	684	GAATTCGTAAAGCGGACGATTGATATGATATGATATGATATGATATGATATGATATGATAT	743	
Db	8426	GAGTTTATCCGTGGTATGTCATGCTGCGCGGAATATATGCTCATCGACCAATAC	8367	
Qy	744	GGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTATTTATAAGAAAAACAGCAA---	799	
Db	8366	GGCATCGGCTACGAGATCTTTTACCGCGAACCGTTTGGGTTTCAGAAAGCGCGAGCGG	8307	
Qy	800	--AGAAACAATCTATACATACATTTATGTAAGAGAGACACGAATGCGCTGTACGGCTTT	857	
Db	8306	CGCGTGACCGTCTATACATACGAATATGACGCGAAGGTGTGACGCGCTTGTACGGGTT	8247	
Qy	858	TCGACAGAGGAGAAAAATCTCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCCA	917	
Db	8246	CCAAACGCGGAGGAGCGCAACTTGTTCGCCAAGCTGTGCAAGTGTTCGGGATCGGGCG	8187	
Qy	918	AAAGAGCGCTTGGCATCTCGCTTCGGCGCATCCGGAGCGGTGATTGAAGCGATCGAG	977	
Db	8186	AAAGCGGCTCGCATTTTGGCGCGCGGCGTTCGATGAATCTGGCGCGGCGCATCGAG	8127	
Qy	978	AATGAGACGAGCATTTCTCGTCAAAATTTCCCGCGGTAGCCAAACCGCAAGCGAG	1037	
Db	8126	GAAGAAAAACGAGCGGTTTATGCAAGTTTCCAGCGTGGGCAAAAGACGCGCCGCA	8067	
Qy	1038	ATCATCTTGCACCTGAAA	1055	
Db	8066	ATGATTGGATTAAAA	8049	
RESULT 11				
AX433201	AX433201	Sequence 1616 from Patent WO0229113.	546 bp	DNA linear PAT 28-JUN-2002
DEFINITION	AX433201	Accession		
ACCESSION	AX433201	Version		
VERSION	AX433201.1	GI:21658005		
KEYWORDS				
SOURCE	Bacillus licheniformis			
ORGANISM	Bacillus licheniformis			
REFERENCE	1	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
AUTHORS	Berka, R. and Clausen, I. G.			
TITLE	Methods for monitoring multiple gene expression			
JOURNAL	Patent: WO 0229113-A 1616 11-APR-2002;			
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)				
FEATURES	Location/Qualifiers			
source	1..546			
	/organism="Bacillus licheniformis"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:1402"			
ORIGIN				
Query Match	18.5%;	Score 195.4;	DB 6;	Length 546;
Best Local Similarity	85.5%;	Pred. No. 2.6e-42;		
Matches	253;	Conservative	0;	Mismatches 37; Indels 6; Gaps 3;
Qy	766	CGCAAAATCGGTTTATTTATAAGAAAAACAGCAAGAAACAATCTATACATA---CCATT	822	
Db	27	CCAGAACCGGTTTATTTATAAGCAACCGCCCAAGCAACATTTCTTATCCATCCATT	86	
Qy	823	ATGTAAGAGAGACACGAATCGCTGTACGCTTTTCGACAGGAGAG-AAAAATGCTG	881	
Db	87	AGGAAGAGAGAGCCCGAATGCGCTGTACGCTTTTCGCCAAGGAGAGAAAAATGCGG	146	

```
Qy 882 TTTACGAAATGCTGAATGTTTACGGGATCGGCCCAAAAGGAGCGCTTGGATCCTCGCT 941
Db 147 TTTTACGAAATGCTGAATGTTTACGGGATCGGCCCAAAAGGAGCGCTTGGATCCTCGCT 206

Qy 942 TCCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGATGAGGACGAAGCATTTCT--CG 999
Db 207 CCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGATGAGGACGAAGCATTTCTTTCGT 266

Qy 1000 TCAATTTCCCGGTAGGCAAAAAACGCGAAGCGAGATCATCTTGACCTGAAA 1055
Db 267 CAAATTTNCCCGGTAGGCAAAAAACGCGAAGCGAGATCATCTTGACCTGAAA 322

RESULT 12
AE017012/c
LOCUS AE017012 301332 bp DNA linear BCT 16-MAY-2003
DEFINITION Bacillus cereus ATCC 14579 section 15 of 18 of the complete genome.
ACCESSION AE017012 AE016877
VERSION AE017012.1 GI:29897926
KEYWORDS
SOURCE Bacillus cereus ATCC 14579
ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 301332)
AUTHORS Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapratral,V., Bhattacharya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goltzman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
TITLE Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis
JOURNAL Nature 423 (6935), 87-91 (2003)
PUBMED 12721630
REFERENCE 2 (bases 1 to 301332)
AUTHORS Candelson,B., Gailloux,K., Ehrlich,D.S. and Sorokin,A.
TITLE The number of ribosomal RNA operons in Bacillus cereus
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 301332)
AUTHORS Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapratral,V., Bhattacharya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goltzman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de Vilvert, Jouy en Josas 78352, France
FEATURES
source
1. 301332
/organism="Bacillus cereus ATCC 14579"
/mol_type="genomic DNA"
/strain="ATCC 14579"
/db_xref="ATCC:14579"
/db_xref="taxon:226900"
complement (114..470)
locus tag="BC4288"
complement (114..470)
locus tag="BC4288"
/codon_start=1
/transl_table=1
/product="Membrane-attached cytochrome c550"
/protein_id="AAP11201.1"
/db_xref="GI:29897927"
/transl_table="MKNRNLIPFALIALGIIVMFVSFQGLNKSKEADAKNGKPA
QTAGSKPDIYVQSCTSCHGDQLOQAVGNLQKIGKLSKDBEILSKGKNMPANIV
PADOAAKVDWLSSKK"
complement (857..1984)
locus tag="BC4289"
complement (857..1984)
locus tag="BC4289"
/codon_start=1
/transl_table=1
/product="RNA polymerase sigma factor rpoD"

/protein_id="AAP11202.1"
/db_xref="GI:29897928"
/transl_table="MADKPARSQIETEMLTQVKEQLTBLGKRGKGVLTVEEIAERMN
GFIEQSDMEYIEGEGIDLVGNDNEEGNNRQISKTEEFDLNDLSVPGVKI
NDPVRMYLKEIGRVDLLSABEEILATRIEEDDEAKRRLAEANRLVSTAKRYVGR
GMFLDLIQBNMGLIKAVEKFDYRKGFESTVATWIRQAIIRAIQARTIRIPVH
MVTINKLIRVQQLQDLGREPSPEIGEMDLAPKPKVLELLKIAQBPVSLPTPIGE
EDSHLGDFTIEDQATSPADHAAYELLKLEOLDVLTLDREENVLRIFGLDDGRTR
TLSEVGKFGVTRERINQIEAKALRKLHRSRSLKLDLFLE"
complement (2044..3840)
locus tag="BC4290"
complement (2044..3840)
locus tag="BC4290"
/EC_number="2.7.7.-"
/codon_start=1
/transl_table=1
/product="DNA primase"
/protein_id="AAP11203.1"
/db_xref="GI:29897929"
/transl_table="MGNRIPEEVVQIOTSSDIVEVIGEVYVQLKQGRNVFGLCPFHG
ENSPFSVSDKQIFHCFCGCGEGNVFSLMKMEGLAPTEAVQKLGKNGIAVAEYTS
GQQQEDISDDTVMQOAHLLKKYHHLLVNTBEGNSALSYLLNRGILTKEMIKFEI
GYASPAWDAATKILQKRLSLTMEQAGLLIRSEKGDHYDRFRGVMPFTYITQGVK
IAPSGRALGDDTPKLYNPETPIFKSKLLYNHQARPFIRKRGQVVLFEQYADVLA
VKSQVEAVATMGALTTEQAKLLRNVTWVLCYDGDKAGREATKAGOLLKLVGCO
VKYTSLPKLDPDPEYQOYCTAFENLVKSSVSPVGKINYLRLGNLQDSSKEEYV
KSVLSELSLQDAQMAESYLSQESFYSMETLLNLHQYRKQKQVQKQIKQVAKP
SQVQTKPLTGFARAREIYIYHMQSAVAVRMESHEDFTEHKGILYELVAYYE
KGNPSVGTFLSWLSDEKLNIIITDIDEFINPEYTEEVQLQGHLETLRHQRKLEKM
EIIPKVKQMEKTPVEAAKYVAYLQNKARK"
4007..4213
locus tag="BC4291"
4007..4213
locus tag="BC4291"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAP11204.1"
/db_xref="GI:29897930"
/transl_table="MLYKIVYICSSVTNLSRNNYTPFYKKGSKPHILFDLEICNFF
MLHITINSILFSIYKDPIALFP"
complement (4202..4333)
locus tag="BC4292"
complement (4202..4333)
locus tag="BC4292"
/codon_start=1
/transl_table=1
/product="EMG2 protein"
/protein_id="AAP11205.1"
/db_xref="GI:29897931"
/transl_table="MQNINILVGVYVPYKAKIVQVGAEPHVGLFVASYAHLPRK"
complement (4701..5513)
locus tag="BC4293"
complement (4701..5513)
locus tag="BC4293"
/codon_start=1
/transl_table=1
/product="hypothetical Cytosolic Protein"
/protein_id="AAP11206.1"
/db_xref="GI:29897932"
/transl_table="MDNKKIVVYSDSVGETADLVVRAAMQGFEPADIRRVYVEDTG
TLKEVISTAKNQALICFTLVKPDROYLATEAKEGVEAYDIIIGPLIDJEETIGQV
PRYEPGVVRRLDEEYFKIEAIEPAVKYDGRDGRGLIKADIVLIGVSRSTKTLPSQY
LAHNKRVVANVPLVPEPPELYQVAKRGKGLKIPDKLNHIRKERLSKSLGSDG
ATTANINRIEEDIDHFEVIRINQVIDVSNKAIBETANIIVNAVQNMKF"
complement (5543..6175)
locus tag="BC4294"
complement (5543..6175)
locus tag="BC4294"
/codon_start=1
/transl_table=1
/product="CBS domain containing protein"
/protein_id="AAP11207.1"
```

```
/db_xref="GI:29897933"
/translation="MIITELNRQBEHIIQIVKDHGPIITCESIAAQLGLTRATLRPLA
ILTMAGYLEARPRVGYFTGKQGLLSEAVKKIKVQDYQSRPVVIDKNVSYDAICT
MFLSDVGLTFVVDQMTLLVGVSRKDLRLASLGKDLTSLPWNIMTRMPNIMCRRE
DSLYDIAMELIERQIDAMPVVKDTKQGLEVIGRIITKNTITRAFVNLVNE"
complement(6283..7029)
/locus_tag="BC4295"
complement(6283..7029)
/locus_tag="BC4295"
/codon_start=1
/transl_table=11
/product="DNA repair protein recO"
/protein_id="AAP11208.1"
/db_xref="GI:29897934"
/translation="MFQKVEGIVRTTDDYGETNKNITVPSRELKGVSAMARGSKPKS
RLKSVQMLTHGHFIQMGSLGTQQGHEIISTMKEIREIDPLTAYASFIVELTDTAK
EDKHNLYLFEMLYOTLHYMCEGVDPEVLSLIYQTKMLPVLGMRPVFTDTCAICHOETD
FVAFVSRGGFLCSRSHSQDPIYRIPVGEAVHKLRLFFHFDLHRLGNVSRKSTKKQM
RLVMTYDVEYCGIYKSRFLEQLDKFQI"
complement(7063..7206)
/locus_tag="BC4296"
complement(7063..7206)
/locus_tag="BC4296"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP11209.1"
/db_xref="GI:29897935"
/translation="MLDTWKFFSGTSETVLLKMEKQVNDQDHEBELAHLDS
FIS"
complement(7383..8288)
/locus_tag="BC4297"
complement(7383..8288)
/locus_tag="BC4297"
/codon_start=1
/transl_table=11
/product="GTP-binding protein"
/protein_id="AAP11210.1"
/db_xref="GI:29897936"
/translation="MNRKGYKGSVSIIGRPNVKSTFLNRIIGQIAIMSDKPTOTR
NKIQGVYENDAQVIFIDTGHKPKHKGDPVKKMAOTLTKEVDIVLPMVNAVGEFG
RGEETILEKLETKQPVFLVINKIDQLHPEQLLELIDQYRKLHFEAEIPIVPSALDQNN
VDALGTIKKIDPEGPYPENQVTDHPERFIIABLRKVLHTRFSPVHSVAVID
AIQKREGAVINATIVVERASQGIIGKQKMLKEVGRARFDIEALLGSKVPLEV
WVKQKDRNKMSQLRDIGFREDEY"
complement(8281..8679)
/locus_tag="BC4298"
complement(8281..8679)
/locus_tag="BC4298"
/EC_number="3.5.4.5"
/codon_start=1
/transl_table=11
/product="Cytidine deaminase"
/protein_id="AAP11211.1"
/db_xref="GI:29897937"
/translation="MNSQLILEAEARQKQAVPYPSKFQVGAALTRGGKVVYRGCNVE
NASYGLNCNAERTALFKAVSEGDFEVAIVVADTKRPVPCGACRQVWVLCQKDTK
VYLSLHGDVQETTVGELLPGAFLEDLHE"
complement(8780..9133)
/locus_tag="BC4299"
complement(8780..9133)
/locus_tag="BC4299"
/EC_number="2.7.1.107"
/codon_start=1
/transl_table=11
/product="diacylglycerol kinase"
/protein_id="AAP11212.1"
/db_xref="GI:29897938"
/translation="MKKGKLIDFQYAIAGIFFCFRHERNMKIHVLAIVVICGFFYF
15.7%; Score 166; DB 1; Length 301332;
Best Local Similarity 62.8%; Pred. No. 5e-34;
Matches 275; Conservative 0; Mismatches 160; Indels 3; Gaps 1;
```

```
QY 621 TTCTAGCTTTGCTGTTTGTGTTTCAATGAAGACAGCTCAAGAGGTGAATGAACGTTG 680
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124209 TTCTATCATTAATATGTTAAATGGAATACATGATTTTAAGAGGAGAGATCGTATTTTG 124150
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 681 ATCGAATTCGTAAGAGGACGATTCGATTTATGTCGCCCAATATATTTGTCATTTGAAAC 740
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124149 TTTGAATATGTTACAGGTTTACGTGGAGTATGTAGGCGCGGAATATGTCGTAATTGATCAT 124090
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 741 GSGGGATCGGCTATCAGATCTTACGCAATCCGTTTATTTATTAAGAAAAACAGCAA 800
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124089 AATGGAATTTGTTATCAAAATTTTACACCAATCCGTATGTTATTTCAAGAAGATGACAA 124030
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 801 GAAA---CAATCTATACATACCATTTATGTAAGAGAAGACACGAATTCGCTGTACGGCTTT 857
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124029 GAAATCCGTGCTATACATATCATTTATGTGAGAGAAGATATTTATGGCACTTTTACGGGTTT 123970
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 858 TCGCAACGGGAAGAAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGATTCGCCCCA 917
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123969 AAAACACGTGAAGAGCGTTTATTTATTTACAAGTTTGTAGGCGTGTCTGGCAATTGGCCA 123910
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 918 AAAGAGCGCTTTCGATCTCGCTTCGCGCATCCGGAGCGGTGATTGAACGATCGAG 977
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123909 AAAGAGCTCTTGCATTTTACATTTTCTGTCACACGACAGGTCGTTCAACGCGATTGAG 123850
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 978 AATGAGGACGAAGCATTTTCTGTCAAATTTCCCGCGGTAGGCAAAAAACGGCAAGGCAG 1037
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123849 CATGAAGATGAGAAGTTTATTAGTGAATTTCCAGGCGTCGGAAGAAACACGACGCCAA 123790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1038 ATCATCTTGCACCTGAAA 1055
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123789 ATGATTTAGATTTAAAA 123772
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
AE017278/c
LOCUS
DEFINITION AE017278 288814 bp DNA linear BCT 23-FEB-2004
Bacillus cereus ATCC 10987, section 15 of 18 of the complete
genome.
ACCESSION AE017278 AE017194
VERSION AE017278.1 GI:42739327
KEYWORDS
SOURCE Bacillus cereus ATCC 10987
ORGANISM Bacillus cereus ATCC 10987
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
REFERENCE
1 (bases 1 to 288814)
Rasko,D.A., Ravel,J., Okstad,O.A., Helgason,E., Cer,R.Z., Jiang,L.,
Shores,K.A., Fouts,D.E., Tourasse,N.J., Angiuoli,S.V., Kolonay,J.,
Nelson,W.C., Kolsto,A.-B., Fraser,C.M. and Read,T.D.
The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
adaptations and a large plasmid related to Bacillus anthracis pX01
Nucleic Acids Res. 32 (3), 977-988 (2004)
JOURNAL
PUBMED 14960714
REFERENCE
2 (bases 1 to 288814)
Rasko,D.A., Ravel,J., Okstad,O.A., Helgason,E., Cer,R.Z., Jiang,L.,
Shores,K.A., Fouts,D.E., Tourasse,N.J., Angiuoli,S.V., Kolonay,J.,
Nelson,W.C., Kolsto,A.-B., Fraser,C.M. and Read,T.D.
Direct Submission
Submitted (19-FEB-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1..288814
/organism="Bacillus cereus ATCC 10987"
/mol_type="genomic DNA"
/strain="ATCC 10987"
/db_xref="ATCC:10987"
/db_xref="taxon:22523"
complement(192..2312)
/locus_tag="BCE4353"
complement(192..2312)
/locus_tag="BCE4353"
/note="Identified by match to protein family HMM PF00905;
```


putative"
/protein_id="AAS43262.1"
/db_xref="GI:42739336"
/translation="MNNILEIEGLSFYEDRNVLIEDINLQVPKGAFLGLVGPNGSGKS
TLKCLGLVLPKQSIIRLFGVDSKKPKEMNKGVYSOKANSFNSGRPATVFEVVSMS
LVSKGLPRFTKNDKKVKRAIDVGNSEFGNIGELSGQQORVFIAIALVSDPE
LLILDPEVGVSDVRNVSPEYILEDINRKLIGITLLVTHDMGAVTEKVTIRVACLNOHL
HFHGNVEKFRLEDAEMSVLYGHVHRLEHEHHGRI"
/complement(9713..10591)
/locus_tag="BCE4362"
/locus_tag="BCE4362"
/locus_tag="BCE4362"
/note="Identified by match to protein family HMM PF02588"
/codon_start=1
/transl_table=11
/product="membrane protein, putative"
/protein_id="AAS43263.1"

Query Match 15.6%; Score 164.4; DB 1; Length 288814;
Best Local Similarity 62.6%; Pred. No. 1.4e-33;
Matches 274; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

Qy 621 TTCTAGCTTTGCTGTTTGTGTTTCAATGAAGACAGTCAAGAGGTTGAATGAACGTTG 680
Db 138093 TTCTATCATTAATATGTTAAATGAATACATGATTTAAGAGGAGAGATTTGATTTTG 138034

Qy 681 ATCGAATTCGTAAGAGGACGATGATGATGATCGCCCAATATATTGTCATTTGAAAC 740
Db 138033 TTTGAATATGTTACAGGTTACGTGAGTATGTAGGACCGGAATATGTCGTAATTGATCAT 137974

Qy 741 GCGGGATCGGTATCAGATCTTACGCGCAATCGGTTTATTTAAGAAACACGCAA 800
Db 137973 AATGGAATTTGGTTATCAAAATTTTACACCGCAATCGGTATGATTTTCAAGAAAGTAAGCAA 137914

Qy 801 GAAA---CAATCTATACATACATTTGTAAGAGACACGATGCGCTGTAGCGCTTT 857
Db 137913 GAATCCGCTCTATACATATCATTTGTAAGAGAGATATTTAGCGGCTTTACGCGTTT 137854

Qy 858 TCGCAAGAGGAGAAATAATGCTTTACGAAATGCTGAATGTTTACGGGGATCGGCCA 917
Db 137853 AAAACACGTGAAGACGTTTATTTATTTACAAAGTTGTTAGTGCTGTTGATTTGGACCA 137794

Qy 918 AAAGAGCGCTTGCATCCTGCTCCGCGATTCGGGAGCGGTGATTTGAAGCGATCGAG 977
Db 137793 AAAGCGCTCTTGCATTTTCTGCTCAATTTCCGCGTAGGCNAAACACGCAAGCGAG 137734

Qy 978 AATGAGGACGAGCAATTTCTGCTCAATTTCCGCGTAGGCNAAACACGCAAGCGAG 1037
Db 137733 CATGAAGACGAGAAAGTTTTTATGTAAGTTCCGCGCGTTCGGAAGAAACACGCGCAA 137674

Qy 1038 ATCATCCTTGACCTGAAA 1055
Db 137673 ATGATTTTAGATTTAAAA 137656

RESULT 14
AE017225_42/c
WPCOMMENT

Sequence split into 53 fragments LOCUS AE017225 Accession AE017225

Fragment Name	Begin	End
AE017225_00	1	110000
AE017225_01	100001	210000
AE017225_02	200001	310000
AE017225_03	300001	410000
AE017225_04	400001	510000
AE017225_05	500001	610000
AE017225_06	600001	710000
AE017225_07	700001	810000
AE017225_08	800001	910000
AE017225_09	900001	1010000
AE017225_10	1000001	1110000
AE017225_11	1100001	1210000
AE017225_12	1200001	1310000

AE017225_13	1300001	1410000
AE017225_14	1400001	1510000
AE017225_15	1500001	1610000
AE017225_16	1600001	1710000
AE017225_17	1700001	1810000
AE017225_18	1800001	1910000
AE017225_19	1900001	2010000
AE017225_20	2000001	2110000
AE017225_21	2100001	2210000
AE017225_22	2200001	2310000
AE017225_23	2300001	2410000
AE017225_24	2400001	2510000
AE017225_25	2500001	2610000
AE017225_26	2600001	2710000
AE017225_27	2700001	2810000
AE017225_28	2800001	2910000
AE017225_29	2900001	3010000
AE017225_30	3000001	3110000
AE017225_31	3100001	3210000
AE017225_32	3200001	3310000
AE017225_33	3300001	3410000
AE017225_34	3400001	3510000
AE017225_35	3500001	3610000
AE017225_36	3600001	3710000
AE017225_37	3700001	3810000
AE017225_38	3800001	3910000
AE017225_39	3900001	4010000
AE017225_40	4000001	4110000
AE017225_41	4100001	4210000
AE017225_42	4200001	4310000
AE017225_43	4300001	4410000
AE017225_44	4400001	4510000
AE017225_45	4500001	4610000
AE017225_46	4600001	4710000
AE017225_47	4700001	4810000
AE017225_48	4800001	4910000
AE017225_49	4900001	5010000
AE017225_50	5000001	5110000
AE017225_51	5100001	5210000
AE017225_52	5200001	5288663

Continuation (43 of 53) of AE017225 from base 4200001 (AE017225 Bacillus anthracis str.

Query Match	15.3%;	Score 161.2;	DB 1;	Length 110000;
Best Local Similarity	62.1%;	Pred. No. 9.7e-33;		
Matches 272;	Conservative 0;	Mismatches 163;	Indels 3;	Gaps 1;

Qy	521	TTCTAGCTTTCCCTGTTTGTGTTTACATGAAGAGCAGTCAAGAGGTTGAATGAACGTTG	680
Db	28998	TTCTATCATTAATATGTTAAATGGAATACATGATTTAAGAGGGAGAGATTTGATTTTG	28839
Qy	681	ATCGAATTCGTAAGAGGACGATTTGATTTATGATCGCCCAATATATTGTCATTTGAAAC	740
Db	28938	TTTGAATATGTTACAGGTTAGTGTAGAGCGGAATATGCTGTAATGATCAT	28779
Qy	741	GCGGGATCGGTATCAGATCTTCAACGCAATTCGTTTATTTATTAAGAAACACGCAAA	800
Db	28778	AATGGAATTTGGTTATCAAAATTTTCAACCGAATTCGTTATTTTCAAGAAAGTAAGCAA	28719
Qy	801	GAAA---CAATCTATACATACATTTGTAAGAGACACGATGCGCTGTACGGCTTT	857
Db	28718	GAAATCCGCTGTATACATATCATTTATGTAAGAGAGATATTTATGGCCATTTTACGGGTTT	28659
Qy	858	TCGCAAGAGGAGAAATAATGCTGTTTACGAAATGCTGAATGTTTACGGGATCGGCCA	917
Db	28658	AAAACACGTGAAGAGCGTTTATTTATTTACAAAGTTGTTAGGGGTGCTGTTGATTTGACCA	28599
Qy	918	AAAGAGCGCTTGCATCCTGCTTCCGCGCATCCGGAGCGGTGATTTGAAGCGATCGAG	977
Db	28598	AAAGGTGCTCTTGCAATTTTAGCTTCTGGTCAACAGCAGGTCGTTTCAAGCGATTGAA	28539
Qy	978	AATGAGGACGAGCAATTTCTGTCATAATTTCCGCGCTAGGCNAAACACGCAAGGAG	1037
Db	28538	CATGAAGACGAGAAAGTTTTTATGTAAGTTCCGCGCGTTCGGAAGAAACACGACGCCAA	28479

Qy 1038 ATCATCCTTGACCTGAAA 1055
Db 28478 ATGATTTTAGATTTAAAA 28461

RESULT 15
AE017334_42/c
WPCOMMENT

Sequence split into 53 fragments LOCUS AE017334 Accession AE017334

Fragment Name	Begin	End
AE017334_00	1	110000
AE017334_01	100001	210000
AE017334_02	200001	310000
AE017334_03	300001	410000
AE017334_04	400001	510000
AE017334_05	500001	610000
AE017334_06	600001	710000
AE017334_07	700001	810000
AE017334_08	800001	910000
AE017334_09	900001	1010000
AE017334_10	1000001	1110000
AE017334_11	1100001	1210000
AE017334_12	1200001	1310000
AE017334_13	1300001	1410000
AE017334_14	1400001	1510000
AE017334_15	1500001	1610000
AE017334_16	1600001	1710000
AE017334_17	1700001	1810000
AE017334_18	1800001	1910000
AE017334_19	1900001	2010000
AE017334_20	2000001	2110000
AE017334_21	2100001	2210000
AE017334_22	2200001	2310000
AE017334_23	2300001	2410000
AE017334_24	2400001	2510000
AE017334_25	2500001	2610000
AE017334_26	2600001	2710000
AE017334_27	2700001	2810000
AE017334_28	2800001	2910000
AE017334_29	2900001	3010000
AE017334_30	3000001	3110000
AE017334_31	3100001	3210000
AE017334_32	3200001	3310000
AE017334_33	3300001	3410000
AE017334_34	3400001	3510000
AE017334_35	3500001	3610000
AE017334_36	3600001	3710000
AE017334_37	3700001	3810000
AE017334_38	3800001	3910000
AE017334_39	3900001	4010000
AE017334_40	4000001	4110000
AE017334_41	4100001	4210000
AE017334_42	4200001	4310000
AE017334_43	4300001	4410000
AE017334_44	4400001	4510000
AE017334_45	4500001	4610000
AE017334_46	4600001	4710000
AE017334_47	4700001	4810000
AE017334_48	4800001	4910000
AE017334_49	4900001	5010000
AE017334_50	5000001	5110000
AE017334_51	5100001	5210000
AE017334_52	5200001	5227419

Continuation (43 of 53) of AE017334 from base 4200001 (AE017334 Bacillus anthracis str.

Query Match 15.3%; Score 161.2; DB 1; Length 110000;
Best Local Similarity 62.1%; Pred. No. 9.7e-33;
Matches 272; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

Qy 621 TTCTAGCTTTGCTGCTTTTGTGTTACAATGAAGAGCAGTCAAGAGGTGAATGAACGTTG 680
Db 28526 TTCTATCATTAATATGTTTAAATGAATACATGATTTTAAGAGGGAGAGATTGTATTG 28467

Qy 681 ATCGAATTCGTAAGAGGACGATTGATTATGATCGCCCAATATATGTCATTGAAAAC 740
Db 28466 TTTGAATATGTTACAGGTTACGTGGAGTATGTAGACCGGAATATGCGTAATTGATCAT 28407
Qy 741 GCGGGATCGGCTATCAGATCTTCAACGCCAAATCCGTTTATTTTATAAGAAAAACAGCAAA 800
Db 28406 AATGGAATTGGTTATCAAAATTTTACACCCGATCCGTTATGTTTCAAGAGTAAGCAA 28347
Qy 801 GAAA---CAATCTATACATACCATTTATGTAAGAGAACACACGAATGCGCTGACGGCTTT 857
Db 28346 GAAATCCGTCTATACATATCATTTATGTGAGAGAGATATTATGGCACCTTTACGGGTTT 28287
Qy 858 TCGACAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTACGGGATCGGCCA 917
Db 28286 AAACACGTGAAGAGCGTTTATTATTACAAAGTTGTTAGGGGTGCTCTGGTATTGGACCA 28227
Qy 918 AAAGGAGCGCTTGCGATCCTCGCTTCCGGGAGTCCGGGAGCGGTGATTTGAAGCGATCGAG 977
Db 28226 AAGGTGCTCTTGCAATTTTAGCTTCTGCTCAACAGACAGGTCGTTCAAGCGATTGAA 28167
Qy 978 AATGAGACGAAGCATTTTCTCGTCAAAATTTCCCGGCGTAGGCAAAAAACCGCAAGGAG 1037
Db 28166 CATGAAGACGAGAAGTTTATTAGTGAAGTTCCTCGGCGCTCGGAAAGAAAAACAGCACG 28107
Qy 1038 ATCATCCTTGACCTGAAA 1055
Db 28106 ATGATTTTAGATTTAAAA 28089

Search completed: February 27, 2006, 08:07:00
Job time : 5721 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:32:21 ; Search time 4975 Seconds
(without alignments)
9921.683 Million cell updates/sec

Title: US-10-510-408-1
Perfect score: 1055
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacctgaaa 1055

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	6.6	767	10	CL678407
2	68.2	6.5	524	2	BG455392
3	64	6.1	790	9	BH371812
4	49.2	4.7	707	8	CX137869
5	47	4.5	468	8	CX067472
6	47	4.5	485	8	CX067473
7	45.8	4.3	745	7	CN384972
8	44.6	4.2	1101	10	CNS0039G
9	42.8	4.1	501	5	BQ496577
10	42.8	4.1	519	5	BQ501436
11	42.8	4.1	527	6	C8533944
12	42.4	4.0	684	8	CX146914
13	42.4	4.0	701	8	CX144590
14	42.4	4.0	738	8	CX153248
15	41.2	3.9	882	9	A2674153
16	41	3.9	859	9	A2540067
17	41	3.9	919	9	BH139275
18	40.8	3.9	468	10	AG936518
19	40.8	3.9	468	10	AG938690
20	40.8	3.9	493	10	AG911866
21	40.8	3.9	546	10	AG944693
22	40.8	3.9	566	10	AG939340

C 23	40.8	3.9	585	10	AG910157
C 24	40.8	3.9	672	10	AG920520
C 25	40.4	3.8	1101	10	CNS0100X
C 26	40	3.8	470	1	AJ761112
C 27	39.8	3.8	405	3	BP813221
C 28	39.8	3.8	796	8	CX949675
C 29	39.8	3.8	855	8	DN640592
C 30	39.8	3.8	905	8	DR887954
C 31	39.6	3.8	338	5	BW578460
C 32	39.6	3.8	542	5	BW585582
C 33	39.4	3.7	374	9	BZ357018
C 34	39.4	3.7	828	7	CK090955
C 35	39.2	3.7	386	10	CZ673753
C 36	39.2	3.7	395	3	BP798691
C 37	39.2	3.7	400	2	BE521520
C 38	39.2	3.7	519	3	BP562728
C 39	39.2	3.7	587	7	CO221100
C 40	39.2	3.7	604	8	DR540570
C 41	39.2	3.7	642	3	BP796746
C 42	39.2	3.7	1524	4	CNS0ADNT
C 43	39.2	3.7	1713	4	CNS0ADK8
C 44	38.8	3.7	1101	10	CNS0125V
C 45	38.8	3.7	1350	10	CL019486

ALIGNMENTS

RESULT 1
CL678407
LOCUS
DEFINITION
PRI0122d_B08.2 - PRI0122d_BR (767) Mixed stage foamid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.
CL678407
CL678407.1 GI:50184707
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
1 (bases 1 to 767)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppADB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
14681447
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seg primer: T7
Class: foamid ends.
Location/Qualifiers
1. .767
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage foamid library of P. pacificus var. California"
/note="Vector: pBpifos-5 Fosmid vector"

FEATURES
source
Query Match 6.6%; Score 70; DB 10; Length 767;
Best Local Similarity 54.5%; Pred. No. 2.3e-08;
Matches 140; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

ORIGIN

Query Match 6.6%; Score 70; DB 10; Length 767;
Best Local Similarity 54.5%; Pred. No. 2.3e-08;
Matches 140; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Mon Feb 27 08:33:23 2006

```

QY 799 AAGAAACAAATCTATACATACCATTTATGTAAGAGAAAGACAGATGCGCTGTACGGCTTTT 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 AAGCGATCGTTTTTCAACCCACTTTTGTGGTGGGTGAAGACGGCGAGCTGCTGTACGGTTTTA 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 859 CGACAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCAA 918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 ACAATAAACAGAGCGCACGTTGTTCAAGAGTTGATCAAAACCAACGCGGTGCGGCCGA 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 919 AAGAGCGCTTGGCATCTCGCTTCCGGCGATCCGGAGCGGTGATTTGAAGCGATCGAGA 978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 AGTTGGCGCTGGCGATCTCTCCGGAATGTGAGCGCAGCACTTCTGTTAATGCCGTTGAGC 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 979 ATGAGGAGCAAGCAATTTCTGTCAAATTTCCGGCGTAGGCAAAAAACGCAAGGCAGA 1038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 GTGAAGAAGTGGGGCGCTGCTGAACTGCCGGGTATTGGCAAAAAAACCGCCGAACGCT 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1039 TCATCCTTGACCTGAAA 1055
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 TGATTGTTGAATGAAA 762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
BG455392/c
LOCUS
DEFINITION
  BG455392 524 bp mRNA linear EST 19-MAR-2001
  clone NF047E06PL 5', mRNA sequence.
ACCESSION
  BG455392
VERSION
  BG455392.1 GI:13378717
KEYWORDS
  Medicago truncatula (barrel medic)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
  rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
  Medicago.
REFERENCE
  1 (bases 1 to 524)
  Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
  Flores, H.R., Iman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.
  Expressed Sequence Tags from the Samuel Roberts Noble Foundation
  Medicago truncatula phosphate-starved leaf library
  Unpublished (2000)
  Contact: Harrison MJ
  Plant Biology Division
  The Samuel Roberts Noble Foundation
  2510 Sam Noble Parkway, Ardmore, OK 73402, USA
  Tel: 580 221 7325
  Fax: 580 221 7380
  Email: mjharrison@noble.org
  Insert Length: 524 Std Error: 0.00
  Plate: 047 row: E column: 06
  Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES
  source
    1..524
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /db_xref="taxon:3880"
    /clone="NF047E06PL"
    /tissue_type="leaf"
    /dev_stage="trifoliolate"
    /clone_lib="Phosphate starved leaf"
    /note="Vector: Lambda Zap; At the trifoliolate stage, M.
    truncatula plants were transplanted to phosphate-free sand
    and grown for a further 30 days. During this 30 day
    period, the plants were fertilized twice weekly with 1/2
    Hoaglands solution containing only 20uM potassium
    phosphate. RNA was prepared from above ground tissues."
ORIGIN
  Query Match 6.5%; Score 68.2; DB 2; Length 524;
  Best Local Similarity 54.1%; Pred. No. 6.8e-08;
  Matches 139; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 799 AAGAAACAAATCTATACATACCATTTATGTAAGAGAAAGACAGATGCGCTGTACGGCTTTT 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 AAGCGATCGTTTTTCAACCCACTTTTGTGGTGGGTGAAGACGGCGAGCTGCTGTACGGTTTTA 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 859 CGACAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCAA 918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 ACAATAAACAGAGCGCACGTTGTTCAAGAGTTGATCAAAACCAACGCGGTGCGGCCGA 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 919 AAGAGCGCTTGGCATCTCGCTTCCGGCGATCCGGAGCGGTGATTTGAAGCGATCGAGA 978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 AGTTGGCGCTGGCGATCTCTCCGGAATGTGAGCGCAGCACTTCTGTTAATGCCGTTGAGC 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 979 ATGAGGAGCAAGCAATTTCTGTCAAATTTCCGGCGTAGGCAAAAAACGCAAGGCAGA 1038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 GTGAAGAAGTGGGGCGCTGCTGAACTGCCGGGTATTGGCAAAAAAACCGCCGAACGCT 745
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1039 TCATCCTTGACCTGAAA 1055
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 746 TGATTGTTGAATGAAA 762
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
BH371812
LOCUS
DEFINITION
  BH371812 790 bp DNA linear GSS 10-DEC-2001
  AG-ND-162110-TR ND-TAM Anopheles gambiae genomic clone
  AG-ND-162110, genomic survey sequence.
ACCESSION
  BH371812
VERSION
  BH371812.1 GI:17317937
KEYWORDS
  GSS.
  Anopheles gambiae (African malaria mosquito)
ORGANISM
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
  Culicidae; Anophelinae; Anopheles.
REFERENCE
  1 (bases 1 to 790)
  Hong, Y.-S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J.,
  Ren, C., Huff, E.R., Carlile, J.L., Black, K., Zhang, H.-B.,
  Gardner, M.J. and Collins, F.H.
  Construction of a BAC library and generation of BAC end
  sequence-tagged connectors for genome sequencing of the African
  malaria mosquito Anopheles gambiae
  Mol. Genet. Genomics 268 (6), 720-728 (2003)
  12655398
  Other GSSs: AG-ND-162110.TF
  Contact: Brendan J Loftus
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0208
  Fax: 301 838 3543
  Email: b.loftus@tigr.org
  This clone is from an A. gambiae BAC library (ND-TAM) provided by
  F.H. Collins and sequenced by The Institute for Genomic Research
  (TIGR). The BAC library was generated from A. gambiae PEST strain
  DNA. All DNA was extracted from newly hatched first instar larvae
  to minimize the inclusion of DNA from microorganisms that inhabit
  the gut. The DNA is derived from mixed sexes of larvae. The BAC
  library was constructed at Texas A&M University BAC Center
  University, College Station, Texas 77843-2123, USA using a HindIII
  partial digest.
  Seq primer: M13 Rev
  Class: BAC ends.
FEATURES
  source
    1..790
    /organism="Anopheles gambiae"
    /mol_type="genomic DNA"
    /strain="PEST"
    /db_xref="taxon:7165"
    /clone="AG-ND-162110"
    /clone_lib="ND-TAM"
    /note="Vector: pECBAC1; Site_1: HindIII"
ORIGIN
  Query Match 6.1%; Score 64; DB 9; Length 790;

```


Mon Feb 27 08:33:23 2006

RESULT 6
CX067473/c
LOCUS
DEFINITION 1322269 NCCWA 04RT Oncorhynchus mykiss cDNA, mRNA sequence.
ACCESSION CX067473
VERSION CX067473.1 GI:56989039
KEYWORDS EST.
SOURCE
ORGANISM
Oncorhynchus mykiss (rainbow trout)
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 485)
Rexroad, C.E., Rise, M., Koop, B., von Schalburg, K. and Yao, J.
OART gpool, NCCWA/WVU EST Project, Phase II, in collaboration with
GRASP
Unpublished (2004)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: 106 row: F column: 8
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1. .485
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/la5_host="Top10"
/clone_lib="NCCWA 04RT"
/note="Vector: PCR 4-TOPO; This is an early neurogenesis
SSH library created by Mathew L. Rise constructed by
subtracting late neurogenesis (mixed stages: hindbrain
swelling + heart tube with peristalsis) from early
neurogenesis (mixed stages: neural groove + 1/2 epiboly).
Fish were from a domesticated strain (Spring Valley Trout
Farm, Langley, B.C.), courtesy of Bob Devlin, DFO. These
are mostly internal (coding) sequences."

FEATURES
source

1. .485
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/la5_host="Top10"
/clone_lib="NCCWA 04RT"
/note="Vector: PCR 4-TOPO; This is an early neurogenesis
SSH library created by Mathew L. Rise constructed by
subtracting late neurogenesis (mixed stages: hindbrain
swelling + heart tube with peristalsis) from early
neurogenesis (mixed stages: neural groove + 1/2 epiboly).
Fish were from a domesticated strain (Spring Valley Trout
Farm, Langley, B.C.), courtesy of Bob Devlin, DFO. These
are mostly internal (coding) sequences."

ORIGIN

Query Match 4.5%; Score 47; DB 8; Length 485;
Best Local Similarity 55.1%; Pred. No. 0.076;
Matches 92; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 491 AACATGTCATGAAGCGCTAAATCATCCGGAAGCCCAACATCATGTAGAAGATATGAAGA 550
DB 183 AACATTCATGAAGATGCAATAATTAAGAATGAGAATGATATTGGAATGTATGTG 124
QY 551 CATGACGCTGTATGCTTTTTCAGCTGAGAGAGAGCTTTTTCAGCAACATATGTT 610
DB 123 CATATGCAAGTTTCTCTTTTAAACAATGAAGATATTTTGGAGAGGATGTATGTT 64
QY 611 AACTTTTTCATCTAGCTTTTCCTGTTTGTGTACCAATGAAGAGCA 657
DB 63 AAAGATTGCTTTTAAATGCAGCATCAATGTGATTTATAGAAAGTA 17

RESULT 7
CN384972
LOCUS
DEFINITION 745 bp mRNA linear EST 31-AUG-2004
Lycopersicon esculentum cDNA clone LE2TR02F22, mRNA sequence.
ACCESSION CN384972
VERSION CN384972.1 GI:51700286
KEYWORDS EST.
SOURCE
ORGANISM
Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
AUTHORS
TITLE

JOURNAL
COMMENT

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 745)
Wang, C.K., Chen, P.Y., Wang, H.M., Soong, S.C., Chen, S.C. and To, K.Y.
DNA microarray profiling of gene expression during tomato root
development
Unpublished (2004)
Contact: Kin-Ying, To
Crop Plant Improvement Group
Institute of Bioagricultural Sciences, Academia Sinica
128 Academia Rd. Section 2, Taipei, Taiwan 11529
Tel: 886-2-2653-3161
Fax: 886-2-2561-5600
Email: kyto@gate.sinica.edu.tw
Insert Length: 745 Std Error: 0.00
Plate: 02 row: F column: 22
Seq primer: smart2.
Location/Qualifiers
1. .745
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="CL5915"
/db_xref="taxon:4081"
/clone_lib="LE2TR02F22"
/tissue_type="roots"
/dev_stage="1-, 2-, 3-, and 4- month-old"
/lab_host="E.coli BM25.8"
/clone_lib="Tomato CL5915 roots under different
developmental stages"
/note="Vector: pRiplex2; Tomato CL5915 seeds are obtained
from AVRDC. Roots were harvested from plants grown under
different developmental stages: 1-, 2-, 3-, 4-month-old.
Equal aliquots of mRNA of different developmental stages
were mixed and used for cDNA library construction. (Smart
PCR cDNA Library construction kit, Clontech)"

FEATURES
source

1. .745
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="CL5915"
/db_xref="taxon:4081"
/clone_lib="LE2TR02F22"
/tissue_type="roots"
/dev_stage="1-, 2-, 3-, and 4- month-old"
/lab_host="E.coli BM25.8"
/clone_lib="Tomato CL5915 roots under different
developmental stages"
/note="Vector: pRiplex2; Tomato CL5915 seeds are obtained
from AVRDC. Roots were harvested from plants grown under
different developmental stages: 1-, 2-, 3-, 4-month-old.
Equal aliquots of mRNA of different developmental stages
were mixed and used for cDNA library construction. (Smart
PCR cDNA Library construction kit, Clontech)"

ORIGIN

Query Match 4.3%; Score 45.8; DB 7; Length 745;
Best Local Similarity 47.6%; Pred. No. 0.19;
Matches 169; Conservative 0; Mismatches 182; Indels 4; Gaps 1;
QY 537 AGAAGATATGAAGACATGACGCTGTATGCTTTTTCAGCTGCAGACAGAGCTTTT 596
DB 358 AGAAGGTGAATATATTTGGAAGATAATAGTATCGGAAGAGAGAGAGAGAGATTTT 417
QY 597 ACAGACATATCTTAACCTTTTTCATCTAGCTTTCCTCTTTTGTGTTTCAATGAAGAGC 656
DB 418 GTCAAAATTTAATTAATCTTAAGAGAAATAGAAGGATGAAGTATTCACACATTCAGAAC 477
QY 657 AGTCAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGGACGATTTGTTATGATCG 716
DB 478 ATGAGAGA- ---AAGGGTAATAACAAGATTTAGAAATTAGTCAGCATGATGTTTAC 533
QY 717 CCCCAATATATGCTCATTTGAAACCGCGGATCGGCTATCAGATCTTCACGCCAAATCCG 776
DB 534 ATTTAATTTTATTTATGCGATCTGTGGATTTTATTTAGTTTAGTTAGTAGTAGTAG 593
QY 777 TTTTATTATGAAGAAACAGCAAGAAACAATCTATACATACCATGATTATGTAAGAGAGAC 836
DB 594 ATTATTAAATTTACACAAATTAAGAAAGATATATGAAGTATACATATGATGATCTAT 653
QY 837 ACGAATGCGCTGTACGCGCTTTTCGCAAGGAGAGAGAAATGCTGTTTACAGAAA 891
DB 654 TGAATTCATTTGGTGTGCTTGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708

RESULT 8
CN00396/c
LOCUS
DEFINITION 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR08K10 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL063921

```
VERSION      AL063921.1  GI:4941778
KEYWORDS     Drosophila melanogaster (fruit fly)
SOURCE       Drosophila melanogaster
ORGANISM     Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
              collaboration with the Berkeley Drosophila Genome Project (BDGP).
              The BDGP is constructing a physical map of the Drosophila
              melanogaster genome using these BACs. For further information
              please see http://www.fruitfly.org The BDGP Drosophila
              melanogaster BAC library was prepared by Kazutoyo Osoegawa and
              Aaron Mammoler in Pieter de Jong's laboratory in the Department of
              Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
              NY. The library is named RPCI-98 and was constructed by partial
              EcoRI digestion of Drosophila DNA provided by the BDGP from the
              isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
              P1 and EST libraries. A more detailed description of the library
              and how to hybridize individual BAC clones, the entire library, or
              filters for hybridization from the BACPAC Resource Center can be
              found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES     Location/Qualifiers
              source            1..1101
                               /organism="Drosophila melanogaster"
                               /mol_type="genomic DNA"
                               /db_xref="taxon:7227"
                               /clone="BACR0810"
                               /clone_lib="RPCI-98"
                               /note="end : TTT3"
ORIGIN
Query Match      4.2%; Score 44.6; DB 10; Length 1101;
Best Local Similarity 14.5%; Pred. No. 0.46;
Matches 62; Conservative 194; Mismatches 171; Indels 0; Gaps 0;
QY 600 GAACATATGTTAACTTTTTCATCTAGCTTCCCTGTTTGTGTTACATGTAAGACGAGT 659
Db 1058 RWAWDGAGTWTATWTTWWWWATWTDWTKWMMWATAAKTDTATWTTWRTAWRADWAGR 999
QY 660 CAAAGAGGTGAATGAAGTTCGAAATTCGTAAGAGGACGATGATTATGTCGCCC 719
Db 998 DRGAKRDRDAITDAGRRDGRKRDKDKDKDDDDDKGKKKKKAAKAATKAWMD 939
QY 720 CAATATATGTCATTGAAACCGCGGATCGGCTATCAGATCTTCAACGCCAAATCCGTTT 779
Db 938 DWDWDKMDKWDGAKDKKADDDGAGDKDDDKGKADDDTDTGDKDDDKDKWDDWDXAG 879
QY 780 ATTTATAGAAAACAGCAAGAAACAATCTATACATCCATATGTAAGAGAAGACAG 839
Db 878 TWGDTATWAAATDWWGWDADWTTWDAADDDWADDDWADWADWADWADWADWADWAD 819
QY 840 AATGCGCTGACGGCTTTTCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899
Db 818 RDWGRAGRGKRGKRRKRDRATWDRDADDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD 759
QY 900 GTTACGGGATCGGCCCAAGAGCGGCTTGGCATCCTCGCTTCGCGGATCCGGGAGCG 959
Db 758 RWAADRTWDRDDDDDDDRAGTAGRKWRRTKRWRTKRWRTKRWRTKRWRTKRWRTKRW 699
QY 960 GTGATTGAAGCGATCAGAGATGAGGACGAGCATTTCTCGTCAAAATTTCCCGCGTAGCG 1019
Db 698 GADAGKGGKTKRKRKRDRATWDRDADWAAAWTTTDTDDDDWDRKRKRKRKRKRKR 639
QY 1020 AAAAAA 1026
Db 638 TTAAAW 632
```

```
RESULT 9
LOCUS      BQ496577
DEFINITION BQ496577 501 bp mRNA linear EST 31-OCT-2002
ACCESSION EST05806 Pb0001 Paracoccidioides brasiliensis cDNA, mRNA sequence.
VERSION    BQ496577.1 GI:24447391
KEYWORDS   EST.
SOURCE     Paracoccidioides brasiliensis
ORGANISM   Paracoccidioides brasiliensis
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Paracoccidioides.
REFERENCE  1 (bases 1 to 501)
AUTHORS    Goldman,G.H., dos Reis Marques,E., Duarte Ribeiro,D.C., de Souza
            Bernades,L.A., Quiapin,A.C., Vitorelli,P.M., Savoldi,M.,
            Semighini,C.P., de Oliveira,R.C., Nunes,L.R., Travassos,L.R.,
            Puccia,R., Batista,W.L., Ferreira,L.E., Moreira,J.C.,
            Bogossian,A.P., Tekala,F., Nobrega,M.P., Nobrega,F.G. and
            Goldman,M.H.
TITLE      Expressed sequence tag analysis of the human pathogen
            Paracoccidioides brasiliensis yeast phase: identification of
            putative homologues of Candida albicans virulence and pathogenicity
            genes
JOURNAL    Eukaryot. Cell 2 (1), 34-48 (2003)
PUBMED    12582121
COMMENT    Contact: Gustavo Henrique Goldman
            Laboratory of Molecular Biology
            Universidade de Sao Paulo - USP - FCFRP
            Av do Cafe S/N, CEP: 14040-903, Ribeirao Preto - SP, Brazil
            Email: ggoldman@usp.br.
FEATURES   Location/Qualifiers
            source            1..501
                               /organism="Paracoccidioides brasiliensis"
                               /mol_type="mRNA"
                               /db_xref="taxon:121759"
                               /clone_lib="Pb0001"
ORIGIN
Query Match      4.1%; Score 42.8; DB 5; Length 501;
Best Local Similarity 55.3%; Pred. No. 1.2;
Matches 83; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 694 AAGGACCATTCGATTATGTCATCGCCCAATATATGTCATTGTAAGAAACGCGGATCGGCT 753
Db 93 AAAAGAGGAGGATCAATCAGCCCAACACACACACAGCGTCGGCAAGCGGAGACGGA 152
QY 754 ATCAGATCTTCACGCCAAATCCGTTTATTTATAGAAAAACAGCAAAACAATCTATA 813
Db 153 ACCACAATTCATTCGAGATCTGATGAAGATAGGAAACATGCGAGACACATACAATT 212
QY 814 CATACCATTTATGTAAGAGACACAGCATG 843
Db 213 TATATAGATATGCAAAAATATACGCGGTG 242
RESULT 10
LOCUS      BQ501436
DEFINITION BQ501436 519 bp mRNA linear EST 31-OCT-2002
ACCESSION EST10661 Pb0001 Paracoccidioides brasiliensis cDNA, mRNA sequence.
VERSION    BQ501436.1 GI:24456338
KEYWORDS   EST.
SOURCE     Paracoccidioides brasiliensis
ORGANISM   Paracoccidioides brasiliensis
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Onygenales; mitosporic Onygenales; Paracoccidioides.
REFERENCE  1 (bases 1 to 519)
AUTHORS    Goldman,G.H., dos Reis Marques,E., Duarte Ribeiro,D.C., de Souza
            Bernades,L.A., Quiapin,A.C., Vitorelli,P.M., Savoldi,M.,
            Semighini,C.P., de Oliveira,R.C., Nunes,L.R., Travassos,L.R.,
            Puccia,R., Batista,W.L., Ferreira,L.E., Moreira,J.C.,
            Bogossian,A.P., Tekala,F., Nobrega,M.P., Nobrega,F.G. and
```

source

1. 527
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-aak-p-13-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-DUI is a normalized cDNA library containing the following tissue(s): Primary Lung Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCGTGTAGGC.
TAG_SEQ=None found"

ORIGIN

Query Match 4.1%; Score 42.8; DB 6; Length 527;
Best Local Similarity 49.1%; Pred. No. 1.2;
Matches 113; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
QY 825 GTAAGAGAGACAGCAATGCGCTGTACGGCTTTTCGACAAAGGAGAAAATGCTGTTT 884
DB 275 GTGCGGAGGAGCGCCAACTGCTGTTGGCTTCTCTACGCGCGGCGAGCGGAATCTTC 216
QY 885 ACGAAATGCTGATGTTACGGGATCGCCCAAAAGGAGCGCTTGGATCTCTCGCTTCC 944
DB 215 CGCGAGCTGATCAAGATCAGCGCGCGTGGGCGCGCANTTGCCTGGCTTGTGCTGCG 156
QY 945 GCGGATCGGAGCGGTGATTGAACGATCGAAGATGAGGACGAGCAATTTCTGTCAAA 1004
DB 155 TTGAGTTGAGGAGGATTTGGCGGAGCGCTGGCGGCGGAGCACGCGCGCTGTGTCAG 96
QY 1005 TTTCCCGCGGTAGGCAAAAACGCGAGGAGATCATCTTGACCTGAA 1054
DB 95 GTGCGGCGCATCGGCAAGAGACGCGCGAGCGCTCTCTGGAATCAA 46

RESULT 12

CX146914/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CX146914 684 bp mRNA linear EST 03-JAN-2005
1286864 NCCWA 03RT Oncorhynchus mykiss cDNA 3', mRNA sequence.
CX146914
CX146914.1 GI:57003569
EST.
Oncorhynchus mykiss (rainbow trout)
Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 684)
Rexroad, C.E., Rise, M., Koop, B., von Schalburg, K. and Yao, J.
03RT rblb, NCCWA/WVU EST Project, Phase II, in collaboration with GRASP
Unpublished (2004)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Lestown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329.

TITLE
Expressed sequence tag analysis of the human pathogen Paracoccidioides brasiliensis yeast phase: identification of putative homologues of Candida albicans virulence and pathogenicity genes
JOURNAL
Eukaryot. Cell 2 (1), 34-48 (2003)
PUBMED
12582121
COMMENT
Contact: Gustavo Henrique Goldman
Laboratory of Molecular Biology - USP - FCFRP
Universidade de Sao Paulo - USP
Av do Cafe S/N, CEP: 14040-903, Ribeirao Preto - SP, Brazil
Email: ggoldman@usp.br.
Location/Qualifiers

FEATURES

source

1. 519
/organism="Paracoccidioides brasiliensis"
/mol_type="mRNA"
/db_xref="taxon:121759"
/clone_lib="PB0001"

ORIGIN

Query Match 4.1%; Score 42.8; DB 5; Length 519;
Best Local Similarity 55.3%; Pred. No. 1.2;
Matches 83; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 694 AAGGACGATGATTATGTTATCGCCCAATATATTGTCATTGAAACGGCGGATCGGCT 753
DB 145 AAAAGGAGGAGTGAATCAGCCCAACACACACAGCGTTCGGCAAGCGGAGACGGGA 204
QY 754 ATCAGATCTTCAGCCCAATCGTTTATTATTAAGAAACACGAGAAACAATCTATA 813
DB 205 ACCACAATTTCATTCGAGATCTGATGAAGANTAGGAAACATGCGAGAGACATCAATT 264
QY 814 CATACATTATGTAAGAGAGACACCAATG 843
DB 265 TATATAGATATGCAAAAATATATACGGGGTG 294

RESULT 11

CB853944/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CB853944 527 bp mRNA linear EST 22-APR-2003
UI-CF-DUI-aak-p-13-0-UI.a1 UI-CF-DUI Homo sapiens cDNA clone
UI-CF-DUI-aak-p-13-0-UI 3', mRNA sequence.
CB853944
CB853944.1 GI:30044321
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 527)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
8895548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.reagen.com) or from Open Biosystems (www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=No.
Location/Qualifiers

FEATURES

```

Plate: 120 row: M column: 2
Seq primer: GTAATACGACTACTATAGGG.
Location/Qualifiers
1. .684
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/tissue_type="pooled"
/clone_lib="NCCWA 03RT"
/lab_host="DH10B"
/notes="Vector: pBluescript SK+; This is a normalized (Cot
= 5) O. mykiss whole juvenile library created by Matthew
L. Rise from approximately 4 month old O. mykiss
(Tzenaicut Lake strain) obtained from Vancouver Island
Trout Hatchery (Duncan, B.C.)."

ORIGIN
Query Match 4.0%; Score 42.4; DB 8; Length 684;
Best Local Similarity 49.5%; Pred. No. 1.7;
Matches 109; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 177 TGAATTTGAGATAAAGTAGAGCGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAA 236
Db 492 TGGTCTAGATCTAGAGCTACCCAACTTACCCCAAGGAGGAGGATACATATGCAATYAA 433
QY 237 AGGTGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAACAGGTGACGA 296
Db 432 AAGATGACTGGTCATTTGGTCAGAAATATCAGATCAGATTTGTGATGTCAAAATTTGGGACCA 373
QY 297 CATTTCTCCCTCAGCAAAACAAACGGTTATATCGGAGTGACTGAAGATGGCGTGATTTC 356
Db 372 AAAACTTTATCCATCTGACACAGGCTAAATTTCCAAAGTGTACGCTGAAGGCATTATCAT 313
QY 357 GACTTTTTCACGGTCGCCCGGGCATCTTATCAGAACCCATT 396
Db 312 AACACTTTACCATTTCTAGTGTATTGACCTCATATT 273

RESULT 13
CX144590/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 701)
Rexroad, C.E., Rise, M., Koop, B., von Schalburg, K. and Yao, J.
03RT rblb, NCCWA/WVU EST Project, Phase II, in collaboration with
GRASP
Unpublished (2004)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: 117 row: G column: 22
Seq primer: GTAATACGACTACTATAGGG.
Location/Qualifiers
1. .701
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/tissue_type="pooled"
/lab_host="DH10B"

FEATURES
source

```

```

/clone_lib="NCCWA 03RT"
/notes="Vector: pBluescript SK+; This is a normalized (Cot
= 5) O. mykiss whole juvenile library created by Matthew
L. Rise from approximately 4 month old O. mykiss
(Tzenaicut Lake strain) obtained from Vancouver Island
Trout Hatchery (Duncan, B.C.)."

ORIGIN
Query Match 4.0%; Score 42.4; DB 8; Length 701;
Best Local Similarity 49.5%; Pred. No. 1.7;
Matches 109; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 177 TGAATTTGAGATAAAGTAGAGCGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAA 236
Db 414 TGGTCTAGAAATCTAGAGCTACCCAACTTACCCCAAGGAGGAGGATACATATGCAAAATAA 355
QY 237 AGGTGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAACAGGTGACGA 296
Db 354 AAGATGACTGGTCATTTGGTCAGAAATATCAGATCAGATTTGTGATGTCAAAATTTGGGACCA 295
QY 297 CATTTCTCCCTCAGCAAAACAAACGGTTATATCGGAGTGACTGAAGATGGCGTGATTTC 356
Db 294 AAAACTTTATCCATCTGACACAGGCTAAATTTCCAAAGTGTACGCTGAAGGCATTATCAT 235
QY 357 GACTTTTTCACGGTCGCCCGGGCATCTTATCAGAACCCATT 396
Db 234 AACACTTTACCATTTCTAGTGTATTGACCTCATATT 195

RESULT 14
CX153248/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oncorhynchus mykiss (rainbow trout)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 738)
Rexroad, C.E., Rise, M., Koop, B., von Schalburg, K. and Yao, J.
03RT rblb, NCCWA/WVU EST Project, Phase II, in collaboration with
GRASP
Unpublished (2004)
Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 x2129
Fax: 304 725 0351
Email: crexroad@nccwa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: 131 row: B column: 19
Seq primer: GTAATACGACTACTATAGGG.
Location/Qualifiers
1. .738
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="NCCWA 03RT"
/notes="Vector: pBluescript SK+; This is a normalized (Cot
= 5) O. mykiss whole juvenile library created by Matthew
L. Rise from approximately 4 month old O. mykiss
(Tzenaicut Lake strain) obtained from Vancouver Island
Trout Hatchery (Duncan, B.C.)."

ORIGIN
Query Match 4.0%; Score 42.4; DB 8; Length 738;

```

	Best Local Similarity	49.58;	Pred. No. 1.7;	Mismatches 111;	Indels 0;	Gaps 0;
	Matches 109;	Conservative	0;			
Qy	177	TGGNATTTGAGAAATAAAGCTAGAGGCCGCTCGCACACTGCGAAGACTTTTAAAGCTGCTTATAA	236			
Db	492	TGGTCTAGAAATCTAGAGCCTACCCAACTTACCCCAAGGAGGAGATACATATGCAATAA	433			
Qy	237	AGGGTGGCAGCTCATCGATCAGAAAAAGGGGTTTATTCTGTTTCGCAAAACAGGGTGGAACA	296			
Db	432	AAGATGACTGGTCAATCGTTCAGAAATAATCAGATCAGATTTGTGATGTCAAAATTTGGACCA	373			
Qy	297	CATTTCTCCCTCAGCAAAACAAACGGTTTATTCGGAGTCACTCAAGATCGCGTGATTTC	356			
Db	372	AAAACTTCATCCCACTCTGAACAGGCTAAAAATTCCAAGTGCAGCTGAAGGGCATTATCAT	313			
Qy	357	GACTTTTTACGGGTGCCCGGGCATCTTTATCAGAACCCATT	396			
Db	312	AACACITTAACATTTCTAGTGTATTTTGACCTCATATT	273			

RESULT 15
AZ674153
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ674153
ENTM064TR Entamoeba histolytica Sheared DNA
genomic, genomic survey sequence.
882 bp DNA linear GSS 14-DEC-2000
GI:11811299
Entamoeba histolytica
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
(bases 1 to 882)
Loftus,B., Van Aken,S. and Fraser,C.
Determination of clone end sequences from Entamoeba histolytica
HM1:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: Shotgun
High quality sequence start: 17
High quality sequence stop: 647.

ORIGIN

Query Match 3.9%; Score 41.2; DB 9; Length 882;
Best Local Similarity 51.6%; Pred. No. 4;

Matches	94;	Conservative	0;	Mismatches	88;	Indels	0;	Gaps	0;
Qy	509	TAAATCATCCGGAAGCCAAACATCATGTGTAGAAGATATGAAGACATGACGCGTGTATGTCT	568						
Db	543	TAATATATTTGATATTTCGAACCTTCATTAATTTGATTTTATATTCAACTGTGCTATGGCT	602						
Qy	569	TTTTTTCAGCTGCAGACAGAGACTTTTTTAGCGAACATATGTTAACTTTTTTTCATCTAGCT	628						
Db	603	TTTTTGTCTCTCCATTGTATTTCTTAATTCGAATTTTTTTTTTTAGTTTGTACTTTTCAT	662						
Qy	629	TTGCCTGTTTGTGTGTACAAATGAAGAGCAGTCAAAGAGGTGAATGAACGTTGATCGAATT	688						
Db	663	TTCGTTTTTATTACATTTAATGATGACATCAACAGAAATTAATCAAACTGATTTGAATT	722						
Qy	689	CG 690							
Db	723	TG 724							

Search completed: February 27, 2006, 07:55:28
Job time : 4979 secs

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:31:46 ; Search time 745 Seconds

(without alignments)
9437.920 Million cell updates/sec

Title: US-10-510-408-1

Perfect score: 1055

Sequence: 1 ccttaagtgtaagcaaaaaa.....agatcatccttgacctgaaa 1055

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055	100.0	1055	12	ADN60329
2	1055	100.0	1522	12	ADG32153
3	459	43.5	459	6	ABK75799
4	195.4	18.5	546	6	ABK74325
5	137	13.0	595	6	ABK78681
6	124.8	11.8	110000	6	ABAO3041_15
7	122	11.6	110000	6	ABQ69245_15
8	122	11.6	110000	6	ABQ67195_1
9	108.2	10.3	110000	9	ADB12064_14
10	107.6	10.2	110000	10	ADP77343_05
11	104.8	9.9	654	9	ADB11501
12	104.8	9.9	654	9	ADB11499
13	103.4	9.8	615	6	ABN91317
14	103.4	9.8	615	13	ADSO1069
15	89.8	8.5	609	8	AB242153
16	89.8	8.5	14273	2	AAV52173
17	89.8	8.5	110000	10	ABS56454_00
18	89	8.4	591	10	ABX05796
19	89	8.4	609	8	ACF72471

20	89	8.4	7159	2	AAV74613
21	88.2	8.4	609	13	ADR92479
22	88.2	8.4	609	14	AEA56349
23	87.8	8.3	110000	6	ABA90521_22
24	87.4	8.3	594	4	AAH90735
25	87.4	8.3	594	13	ADK45627
26	87.4	8.3	603	14	ADW94411
27	87.4	8.3	618	4	AAH90848
28	87.4	8.3	11451	14	ADW94161
29	84.8	8.0	2121	6	ABQ70788
30	74.4	7.1	696	10	ADH85304
31	72.2	6.8	594	6	ABN67714
32	71.8	6.8	110000	6	ABN71527_20
33	70.2	6.7	43539	13	ADV87729
34	70.2	6.7	43539	13	ADV78982
35	70.2	6.7	110000	13	ADV81204_21
36	70	6.6	588	6	ABN67713
37	68.4	6.5	591	13	ADV85324
38	65.6	6.2	633	11	ABD01183
39	65.6	6.2	10996	2	AAK13139
40	65.6	6.2	10996	6	ABS98934
41	61.4	5.8	624	12	ADL02869
42	61.4	5.8	45613	4	AAF28535
43	60.2	5.7	609	9	ADA30941
44	60.2	5.7	657	10	ADF00076
45	56	5.3	618	10	ACF68290

ALIGNMENTS

RESULT 1

ADN60329

ID ADN60329 standard; DNA; 1055 BP.

XX

AC ADN60329;

XX

DT 01-JUL-2004 (first entry)

XX

DE B. licheniformis sporulation related polynucleotide, seq id 1.

XX

KW Mutant host cell; sporulation; oxidoreductase; transferase; hydrolase;

KW lyase; isomerase; ligase; gene; ds.

XX

OS Bacillus licheniformis.

XX

PN WO2003087148-A2.

XX

PD 23-OCT-2003.

XX

PF 25-MAR-2003; 2003WO-DK000200.

XX

PR 10-APR-2002; 2002DK-00000533.

XX

PA (NOVO) NOVOZYMES AS.

XX

PI Andersen JT, Jorgensen ST, Rasmussen MD, Olsen PB, Clausen IG;

XX

DR WPI; 2004-122131/12.

XX

PT P-PSDB; ADN60330.

XX

PT A Bacillus licheniformis mutant host cell for producing a product of

XX

PS interest e.g. vitamins, antibiotics and enzymes.

XX

PS Disclosure; SEQ ID NO 1; 319pp; English.

XX

CC The invention relates to a Bacillus licheniformis mutant host cell derived from a parent B. licheniformis host cell. The mutant host cell is mutated in one or more genes encoding one or more polypeptides involved in sporulation. The host cell comprises one or more heterologous genes present in at least two copies, encoding one or more heterologous polypeptides. The heterologous genes are stably integrated into the genome of the cell without leaving any antibiotic resistance marker genes

at the site of integration. The heterologous genes are transcribed from a heterologous promoter or from an artificial promoter. The heterologous operon, preferably a polyclonal operon. The heterologous peptide is an antimicrobial peptide, or a fusion peptide comprising a peptide part which in its native form has antimicrobial activity. The heterologous polypeptide is an enzyme, preferably a secreted enzyme. The enzyme is an enzyme of a class selected from the group of enzyme classes consisting of oxidoreductases (EC 1), transferases (EC 2), hydrolases (EC 3), lyases (EC 4), isomerases (EC 5), and ligases (EC 6). The Bacillus licheniformis is useful in a process for producing at least one product of interest, comprising cultivating a B. licheniformis mutant host cell in a suitable medium, whereby the said product is produced. The process further comprises isolating or purifying the product of interest. The current sequence represents a B. licheniformis sporulation related polynucleotide.

Sequence 1055 BP; 325 A; 208 C; 259 G; 263 T; 0 U; 0 Other;

Query Match 100.0%; Score 1055; DB 12; Length 1055;
Best Local Similarity 100.0%; Pred. No. 9.8e-314;
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTTAAGTAAGGCAAAAAGAGGTGATATGTATGATGACGCGCAAGCAAGTTCAAAATC 60
DB 1 CCTTAAGGTAAGGCAAAAAGAGGTGATATGTATGATGACGCGCAAGCAAGTTCAAAATC 60
QY 61 GGTATTATGCTTATGGAAGTCTGTCGGCGGCTCAGCTTTCACCTGGAGGCTTGCC 120
DB 61 GGTATTATGCTTATGGAAGTCTGTCGGCGGCTCAGCTTTCACCTGGAGGCTTGCC 120
QY 121 GAAAAGCGGCTTAAAGTTCAATCAGCTTGAAGAGGTTTATCTGGACGAGAGCTTGGGA 180
DB 121 GAAAAGCGGCTTAAAGTTCAATCAGCTTGAAGAGGTTTATCTGGACGAGAGCTTGGGA 180
QY 181 ATTGAGAATAAAGTAGAGCCGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAAGGG 240
DB 181 ATTGAGAATAAAGTAGAGCCGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAAGGG 240
QY 241 TGGGAGCTTCATGATCAGAAAAGGGTTTATCTGTTTCGCAACAGCTGACGACATTT 300
DB 241 TGGGAGCTTCATGATCAGAAAAGGGTTTATCTGTTTCGCAACAGCTGACGACATTT 300
QY 301 TCTCCCTCAGCAAAAACAAAGGTTTATCTGAGTGAAGTGGGCTGATTTCCGACT 360
DB 301 TCTCCCTCAGCAAAAACAAAGGTTTATCTGAGTGAAGTGGGCTGATTTCCGACT 360
QY 361 TTTCAGGTCGCGCGGCTATTTATCAGAACCCATTCATCGTTTTCAGATTGATATA 420
DB 361 TTTCAGGTCGCGCGGCTATTTATCAGAACCCATTCATCGTTTTCAGATTGATATA 420
QY 421 AAGCGGCTGGAAGCCGATGCGGATGATCTGCGCAAGGATACCATACCGACGAAA 480
DB 421 AAGCGGCTGGAAGCCGATGCGGATGATCTGCGCAAGGATACCATACCGACGAAA 480
QY 481 AAGGAATTGAACATGTATTGAAGCCGTAATCATCGGAAGCAACATCATGTAGAA 540
DB 481 AAGGAATTGAACATGTATTGAAGCCGTAATCATCGGAAGCAACATCATGTAGAA 540
QY 541 GATATGAAGACATGACGCTGTTTATGCTTTTTCAGCTGCGACAGAGCTTTTTAGCG 600
DB 541 GATATGAAGACATGACGCTGTTTATGCTTTTTCAGCTGCGACAGAGCTTTTTAGCG 600
QY 601 AACATATGTTAATCTTTTTCATGCTTGTGCTTTTGTGTTACATGAAGACGATC 660
DB 601 AACATATGTTAATCTTTTTCATGCTTGTGCTTTTGTGTTACATGAAGACGATC 660
QY 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATTGATTATGATCGCCCC 720
DB 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATTGATTATGATCGCCCC 720
QY 721 AATATATGTTCAATGAAAACGCGGATCGGCTATCAGATCTTCACGCCAAATCGGTTTA 780
DB 721 AATATATGTTCAATGAAAACGCGGATCGGCTATCAGATCTTCACGCCAAATCGGTTTA 780

QY 781 TTTATAAGAAAAACAGCAAAAGAAACAATCTATACATACCATTTATGTAGAGAACACAGA 840
DB 781 TTTATAAGAAAAACAGCAAAAGAAACAATCTATACATACCATTTATGTAGAGAACACAGA 840
QY 841 ATGCGCTGTACGGCTTTTCGACAAAGGAAGAAAAAATGCTGTTTACGAAAAATGCTGAATG 900
DB 841 ATGCGCTGTACGGCTTTTCGACAAAGGAAGAAAAAATGCTGTTTACGAAAAATGCTGAATG 900
QY 901 TTACGGGATCGGCCCCAAAAGAGGCGTTTGGATCCTCGCTTCGGCGATCCGGGAGCGG 960
DB 901 TTACGGGATCGGCCCCAAAAGAGGCGTTTGGATCCTCGCTTCGGCGATCCGGGAGCGG 960
QY 961 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGCTAGGCA 1020
DB 961 TGATTGAAGCGATCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGCTAGGCA 1020
QY 1021 AAAAAACGGCAAGCGAGATCATCTTCACCTGAAA 1055
DB 1021 AAAAAACGGCAAGCGAGATCATCTTCACCTGAAA 1055

RESULT 2
ADG32153

ID ADG32153 standard; DNA; 1522 BP.

XX AC ADG32153;

XX XX

DT 26-FEB-2004 (first entry)

DE DNA encoding a mutant B. licheniformis secreted polypeptide SeqID 123.

XX mutant; host cell; production yield; shelf life; product stability;

XX purity; secreted; gene; db.

OS Synthetic.

OS Bacillus licheniformis.

XX WO2003093453-A2.

XX 13-NOV-2003.

XX 25-MAR-2003; 2003WO-DK000198.

XX 10-APR-2002; 2002DK-00000534.

XX (NOVO) NOVOZYMES AS.

XX Andersen JT, Jorgensen ST, Rasmussen MD, Olsen PB, Clausen IG;

XX WPI; 2004-053045/05.

XX P-FSDB; ADG32154.

XX New mutant Bacillus licheniformis host cell secreting 5 % less of one or more secreted polypeptides than the parent host cell, useful for producing a product of interest e.g. polypeptides, amino acids or carbohydrates.

XX Disclosure; SEQ ID NO 123; 422pp; English.

XX This invention relates to a novel Bacillus licheniformis (B. licheniformis) mutant host cell derived from a parent B. licheniformis host cell that is mutated in genes encoding secreted polypeptides.

XX Specifically, it refers to the generation of an improved Bacillus host that reduces the need for product purification caused by contaminant secreted native polypeptides in the culture medium. Accordingly, the present invention describes reducing the expression of these native proteins (e.g. proteolytic enzymes, nutrient uptake factors and signal molecules), which in turn makes it easier to purify the heterologous product of interest and therefore improving the production process.

XX Further benefits of a mutated host cell include an increase in total production yield and a longevity of shelf life attributable to improved product stability and purity. This polynucleotide is a DNA sequence


```
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 65 TATTGCTTATTGGAAGTCTGTGCGCGCTCAGCTTTCACCTGGAGCCCTTGCGCGAAA 124
Db |||||||
Qy 1 TATTGCTTATTGGAAGTCTGTGCGCGCTCAGCTTTCACCTGGAGCCCTTGCGCGAAA 60
Db |||||||
Qy 125 AGCGCGCTTAAAGTTCAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 184
Db |||||||
Qy 61 AGCGCGCTTAAAGTTCAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 120
Db |||||||
Qy 185 AGAATAAAGTAGAGCCGCTCCACACTGGAAGACTTTAAAGCTGCTTTATAAAGGTGGC 244
Db |||||||
Qy 121 AGAATAAAGTAGAGCCGCTCCACACTGGAAGACTTTAAAGCTGCTTTATAAAGGTGGC 180
Db |||||||
Qy 245 AGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTGACGACATTTCCTC 304
Db |||||||
Qy 181 AGCTCATCGATCAGAAAAGGGTTTATCTGTTTCGCAACAGGTGACGACATTTCCTC 240
Db |||||||
Qy 305 CCCTCAGCAAAACAAAGGTTTATCTGGAGTGACTGAAGTGCGTGATTTTCGACTTTTC 364
Db |||||||
Qy 241 CCCTCAGCAAAACAAAGGTTTATCTGGAGTGACTGAAGTGCGTGATTTTCGACTTTTC 300
Db |||||||
Qy 365 ACGETGCGCGCGGATCTTTATCAGAACCCATTCATCGTTTTTTCAGATTGATATAAGC 424
Db |||||||
Qy 301 ACGETGCGCGCGGATCTTTATCAGAACCCATTCATCGTTTTTTCAGATTGATATAAGC 360
Db |||||||
Qy 425 GCCTGGAAGCCGATGCGCGATGATCTGCGCAAGGATACCATACCGCAGCAAAAAGG 484
Db |||||||
Qy 361 GCCTGGAAGCCGATGCGCGATGATCTGCGCAAGGATACCATACCGCAGCAAAAAGG 420
Db |||||||
Qy 485 AATTGGAACATGTCTAATCAAGCCGTAAATCATCCGGAA 523
Db |||||||
Qy 421 AATTGGAACATGTCTAATCAAGCCGTAAATCATCCGGAA 459
Db |||||||

RESULT 4
ABK74325
ID ABK74325 standard; DNA; 546 BP.
XX
XX
AC ABK74325;
XX
XX
13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #1616.
XX
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
XX
PN WO200229113-A2.
XX
XX
PD 11-APR-2002.
XX
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
XX
PR 06-OCT-2000; 2000US-00680598.
XX
XX
PA (NOVO ) NOVOZYMES BIOTECH INC.
PA (NOVO ) NOVOZYMES AS.
XX
XX
PI Berka R, Clausen IG;
XX
XX
DR WPI; 2002-416684/44.
XX
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
XX
PS Claim 4; SEQ ID NO 1616; 200pp; English.
```

```
XX The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC -up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 546 BP; 165 A; 123 C; 141 G; 116 T; 0 U; 1 Other;
Query Match 18.5%; Score 195.4; DB 6; Length 546;
Best Local Similarity 85.5%; Pred. No. 3.7e-49;
Matches 253; Conservative 0; Mismatches 37; Indels 6; Gaps 3;
Qy 766 CGCCAAATCCGTTTATTATAAGAAAACAGCAAGAAAACAATCTATACATA---CCATT 822
Db |||||||
Qy 27 CCCAGAACCCGTTTATTATAAGCAAAACCCGCCAAGCAACATTCTTATCCATCCCATTT 86
Db |||||||
Qy 823 ATGTAAGAAAGACACGAATGCGCTGTACGGCTTTTCGACAGAGGAAG-AAAAAATGCTG 881
Db |||||||
Qy 87 AGGAAAGAAAGAGCCCCGAATGCGCTGTACGGCTTTTCGCCAAGGAGGAAAAAATGCGG 146
Db |||||||
Qy 882 TTTACGAAATGCTGAATGTTACGGGGATCGCCCCAAAAGAGGCGCTTGCGATCCTCGCT 941
Db |||||||
Qy 147 TTTACGAAATGCTGAATGTTACGGGGATCGCCCCAAAAGAGGCGCTTGCGATCCTCGCT 206
Db |||||||
Qy 942 TCCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGAAATGAGGACGAGCATTTCT--CG 999
Db |||||||
Qy 207 CCCGGCGATCCGGAGCGGTGATTGAAGCGATCGAGAAATGAGGACGAGCATTTCTTCTG 266
Db |||||||
Qy 1000 TCAAATTTCCCGGTAGGCAAAAAACGGCAAGGCGAGATCATCTTGACCTGAAA 1055
Db |||||||
Qy 267 CAAATTTNCCCGGTAGGCAAAAAACGGCAAGGCGAGATCATCTTGACCTGAAA 322
Db |||||||

RESULT 5
ABK78681
ID ABK78681 standard; DNA; 595 BP.
XX
XX
AC ABK78681;
XX
XX
13-AUG-2002 (first entry)
XX
DE Bacillus clausii genomic sequence tag (GST) #1524.
XX
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus clausii.
XX
XX
PN WO200229113-A2.
XX
XX
PD 11-APR-2002.
XX
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
XX
PR 06-OCT-2000; 2000US-00680598.
```

27-MAR-2001; 2001US-0279526P.

(NOVO) NOVOZYMES BIOTECH INC.
(NOVO) NOVOZYMES AS.

Berka R, Clausen IG;
WPI; 2002-416684/44.

Monitoring differential expression of several genes in first Bacillus cell relative to expression of same genes in one or more second Bacillus cells, by using substrate containing Bacillus genomic sequenced tag array.

Claim 11; SEQ ID NO 5972; 200pp; English.

The invention describes a method of monitoring differential expression of genes in a first Bacillus cell relative to expression of the genes in other Bacillus cells, comprising hybridising labelled nucleic acid probes isolated from Bacillus cells to a substrate containing array of Bacillus genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first Bacillus cell relative to expression of the same genes in one or more second Bacillus cells. The method is useful for monitoring global expression of several genes from a Bacillus cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which Bacillus cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences

Sequence 595 BP; 183 A; 115 C; 145 G; 152 T; 0 U; 0 Other;

Query Match 13.0%; Score 137; DB 6; Length 595;
Best Local Similarity 61.9%; Pred. No. 3.8e-31;
Matches 237; Conservative 0; Mismatches 140; Indels 6; Gaps 1;

679 TGATCGAATTCGTTAAAGGACGATTTGATTTGATTCGCCCCCAATATATTTGTCATTGAAA 738
4 TGATTGATTATTAAGGGAATTTGGTTGCAGTTGAACCTGCTTATCTCGTGGTCGAGA 63
739 ACGCGGGATCGGCTATCAGATCTTTCACGCCAAATCC-----GTTTATTTATAGAAAA 792
64 CCAACGGGATGGGTTCCAAATTTACTGTGCAACCCCTTACCGTTTTTATAAGCGATTAA 123
793 ACACCAAGAACATCTATACATACCATTTATTAAGAGACACGATCGGCTGTACG 852
124 ATCGGGAGTATGTATCCCGACACACCATCTATTTTCGCGAAGACACGCCCTATTGT 183
853 GCTTTTCACAAAGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCG 912
184 GCTTTTACCAACAGACGACGATCTGTTGTTTGAAGACCTTTGATGTCCTGCGCATG 243
913 CCCCCAAAGGACGCTTCGATCTCGTTCCGGCGATCCGGGACGGGTGATTGAAGCGA 972
244 GCCCCAAAGGACGCTTCGATCTCGTTCCGGGACGAGATATTTATCCATGCAA 303
973 TCGAATGAGACGACGATCTCTGTCAAATTTCCCGGGGTAGGCAAAAAACGGCAA 1032
304 TTGAACAGAGATGAGGACCTGCTGTTGTCGTTTCCCTGGAGTCGGCAAAAAACAGCA 363
1033 GGCAGATCATCTTGCACCTGAAA 1055
364 GACAAATCATTTCTTGACTTAAAA 386

Query Match 13.0%; Score 137; DB 6; Length 595;
Best Local Similarity 61.9%; Pred. No. 3.8e-31;
Matches 237; Conservative 0; Mismatches 140; Indels 6; Gaps 1;

RESULT 6

ABA03041_15/c
Continuation (16 of 30) of ABA03041 from base 1500001 (Listeria monocytogenes EGD-e geno
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041

WP	Fragment Name	Begin	End
WP	ABA03041_00	1	110000
WP	ABA03041_01	100001	210000
WP	ABA03041_02	200001	310000
WP	ABA03041_03	300001	410000
WP	ABA03041_04	400001	510000
WP	ABA03041_05	500001	610000
WP	ABA03041_06	600001	710000
WP	ABA03041_07	700001	810000
WP	ABA03041_08	800001	910000
WP	ABA03041_09	900001	1010000
WP	ABA03041_10	1000001	1110000
WP	ABA03041_11	1100001	1210000
WP	ABA03041_12	1200001	1310000
WP	ABA03041_13	1300001	1410000
WP	ABA03041_14	1400001	1510000
WP	ABA03041_15	1500001	1610000
WP	ABA03041_16	1600001	1710000
WP	ABA03041_17	1700001	1810000
WP	ABA03041_18	1800001	1910000
WP	ABA03041_19	1900001	2010000
WP	ABA03041_20	2000001	2110000
WP	ABA03041_21	2100001	2210000
WP	ABA03041_22	2200001	2310000
WP	ABA03041_23	2300001	2410000
WP	ABA03041_24	2400001	2510000
WP	ABA03041_25	2500001	2610000
WP	ABA03041_26	2600001	2710000
WP	ABA03041_27	2700001	2810000
WP	ABA03041_28	2800001	2910000
WP	ABA03041_29	2900001	2944528

Query Match 11.8%; Score 124.8; DB 6; Length 110000;
Best Local Similarity 58.2%; Pred. No. 3.7e-26;
Matches 241; Conservative 0; Mismatches 167; Indels 6; Gaps 1;

648 ATGAAGACGACTCAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGACGATTGAT 707
69055 ATGACGTGAATGAAGAGGACGCAATAAATTTGTACGATTATACATAAAAGGACCGTTACG 68996
708 TATGTATCGCCCCCAATATATTTGTCATTCGAAAACGCGGATCGGCTATCAGATCTTCACG 767
68995 ACGATTACACCCGGAATATATTTGTTGTGAGGACGACAAATCGGCTATCAATAATATCA 68936
768 CCAATCCGTTTATTTATAAGA-----AAACACGCAAGAGGACCAATCTATACATACCAT 821
68935 GGGAAACCGTTTTCCTTTCAACGACTAGAGGTACAGAGGCGCAAGTCTTTTGTATCAG 68876
822 TATGTAGAGAGACACGAAATCGCTGTACCGCTTTTCGACAGGGAAGAAAAATGCTG 881
68875 CATGTGAGAGAGGATAATATTTCTTTATTTGTTTTCAAACACACAGAGAAGCGTTATTTA 68816
882 TTTACGAAAAATGCTGAATGTTTACGGGATCCGCCCAAGAGCGCTTTCGATCCCTCGT 941
68815 TTCAAAAATTTTGTAGTGTTCGGGCATTCGACCAAAAGCGCACTAGCCATTATTCG 68756
942 TCCGCGCATCCGGGACGGTGTATTTGAACGCGATCGAGATGAGGACGAAAGCAATTTCTCGTC 1001
68755 TCAGCGATGTCGTTCCACTAATTTTCAGCGATTGAATCCGAAGAGATGTTTATTTAACC 68696
1002 AAATTTCCCGCGGTAGGCAAAAAACGCAAGGCGAGATCATCTTTCGACCTGAAA 1055
68695 AAATTTCCGAGTGTGCGTAAAAAACAACGCTCGCCAAATTTATTTCTTGATTTAAAA 68642

RESULT 7

ABQ69245_15/c
Continuation (16 of 31) of ABQ69245 from base 1500001 (Listeria innocua DNA sequence #68
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession ABq69245


```
CC protein from the present invention.
XX
SQ Sequence 654 BP; 192 A; 148 C; 155 G; 159 T; 0 U; 0 Other;

Query Match          9.9%; Score 104.8; DB 9; Length 654;
Best Local Similarity 56.0%; Pred. No. 3.3e-21;
Matches 243; Conservative 0; Mismatches 182; Indels 9; Gaps 2;

QY 628 TTTGCCCTGTTTGTGTACATGAGACAGCTCAAGAGGTGAATGACGTTGATCGAAT 687
Db 2 TTTATCTGTGTAATAAATGACTATTGGAAAAAGTGGTGAAGTA---TGTATGAAT 58

QY 688 TCGTAAAGGACGATGATTATGTATCGCCCCCAATATATTGTCATTGAAAAACGGCGGGA 747
Db 59 ACATGGTAGGACAGCTGCTGTGATGTCACCCAGCTACTTGTGTTTACAAGTGGGGGAG 118

QY 748 TCGGCTATCAGATCTTCACGCCAAATCCGTTTATTTATTAAGAAAA-----ACAGCAAAG 801
Db 119 TGGGCTACCACTTTTAATGGCTAACCCCTTCCGCTGAACGACCGGCTGCGCCCAAGAAG 178

QY 802 AAACAATCTATACATACATATGTAAGAGACAGCAATCGCGCTGTACGGCTTTTCGA 861
Db 179 TGAAGATTTACCTTTACCTGTAATGTTAGTCAAGACCAAGCTTAGCTCTTTGTTTCCCC 238

QY 862 CAAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTTACGGGATCGGCCCAAAAG 921
Db 239 GCCAGATGAAAAAGAGCTTGTTCATAAATTAATCAATGCTCAGGTATTGGGCCCAAAA 298

QY 922 GAGCGCTTGGCATCTTCGCTTCCGGCGATCCGGGAGCGGTGATTAAGCGCATCGAGAATG 981
Db 299 GTGCTTTGGCTATCTCGCCCAATGAAGACCATGCTGGCTTAATCCAGGCGATTGAAAAATG 358

QY 982 AGGACGAGCATTTCTCTCAAAATTTCCGGGTAGGCAAAAAACGCAAGGCAGATCA 1041
Db 359 AAGACCGACAGTTTTTGGCAACGCTTTCCGGGTAGGAAAAAACAAGCTCTCAGATTG 418

QY 1042 TCCTTGACCTGAAA 1055
Db 419 TTCTGGACCTCAAA 432

RESULT 12
ADBI1499
ID ADBI1499 standard; DNA; 654 BP.
XX
AC ADBI1499;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloiococcus otitis antigenic protein encoding DNA SEQ ID NO:5563.
XX
KW Alloiococcus otitidis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection; gene; ds.
XX
OS Alloiococcus otitis.
XX
PN WO2003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX
PR 29-NOV-2001; 2001US-0333777P.
XX
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP ) WYETH HOLDINGS CORP.
XX
PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX
WP1; 2003-505284/47.
XX
DR P-PSDB; ADBI1496.
XX
PT New Alloiococcus otitidis polynucleotides and polypeptides, useful for
```

```
PT treating and diagnosing diseases, drug screening assays and monitoring of
XX effects during drug clinical trials.
XX
PS Claim 7; SEQ ID NO 5563; 1019pp; English.
XX
The present invention describes an isolated polynucleotide (I) of
CC Alloiococcus otitidis genomic DNA, which encodes an antigenic protein.
CC Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
CC expression vector comprising the novel isolated polynucleotide (I); its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiococcus otitidis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Alloiococcus
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (I) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiococcus
CC otitidis. The present sequence encodes an Alloiococcus otitidis antigen
CC protein from the present invention.
XX
SQ Sequence 654 BP; 192 A; 148 C; 155 G; 159 T; 0 U; 0 Other;

Query Match          9.9%; Score 104.8; DB 9; Length 654;
Best Local Similarity 56.0%; Pred. No. 3.3e-21;
Matches 243; Conservative 0; Mismatches 182; Indels 9; Gaps 2;

QY 628 TTTGCCCTGTTTGTGTACATGAGACAGCTCAAGAGGTGAATGACGTTGATCGAAT 687
Db 2 TTTATCTGTGTAATAAATGACTATTGGAAAAAGTGGTGAAGTA---TGTATGAAT 58

QY 688 TCGTAAAGGACGATGATTATGTATCGCCCCCAATATATTGTCATTGAAAAACGGCGGGA 747
Db 59 ACATGGTAGGACAGCTGCTGTGATGTCACCCAGCTACTTGTGTTTACAAGTGGGGGAG 118

QY 748 TCGGCTATCAGATCTTCACGCCAAATCCGTTTATTTATTAAGAAAA-----ACAGCAAAG 801
Db 119 TGGGCTACCACTTTTAATGGCTAACCCCTTCCGCTGAACGACCGGCTGCGCCCAAGAAG 178

QY 802 AAACAATCTATACATACATATGTAAGAGACAGCAATCGCGCTGTACGGCTTTTCGA 861
Db 179 TGAAGATTTACCTTTACCTGTAATGTTAGTCAAGACCAAGCTTAGCTCTTTGTTTCCCC 238

QY 862 CAAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTTACGGGATCGGCCCAAAAG 921
Db 239 GCCAGATGAAAAAGAGCTTGTTCATAAATTAATCAATGCTCAGGTATTGGGCCCAAAA 298

QY 922 GAGCGCTTGGCATCTTCGCTTCCGGCGATCCGGGAGCGGTGATTAAGCGCATCGAGAATG 981
Db 299 GTGCTTTGGCTATCTCGCCCAATGAAGACCATGCTGGCTTAATCCAGGCGATTGAAAAATG 358

QY 982 AGGACGAGCATTTCTCTCAAAATTTCCGGGTAGGCAAAAAACGCAAGGCAGATCA 1041
Db 359 AAGACCGACAGTTTTTGGCAACGCTTTCCGGGTAGGAAAAAACAAGCTCTCAGATTG 418

QY 1042 TCCTTGACCTGAAA 1055
Db 419 TTCTGGACCTCAAA 432

RESULT 13
ABN91317
```



```
SQ Sequence 615 BP; 241 A; 90 C; 110 G; 174 T; 0 U; 0 Other;

Query Match          9.8%; Score 103.4; DB 13; Length 615;
Best Local Similarity 56.4%; Pred. No. 8.6e-21;
Matches 216; Conservative 0; Mismatches 161; Indels 6; Gaps 1;

QY 679 TGATCGAATTCGTAAGGGACGATTGATTATGATCGCCCAATATATATTCATTCATTTGAAA 738
DB 14 TGTATGCATATATTAAAGGAACATATCTCAGTTGTTCCCTACACATGATGTTGGTTGAGA 73

QY 739 ACGGGGGATCGGCTATCAGATCTTTCAGCCAAATCCGTTTATTATTAAGAAAAACAGCA 798
DB 74 CATGTGGCATAGGTTATGAGATACAAAGCCCTAATCTCTATCGTTTTCANAAATATCTTG 133

QY 799 AAGAAACATC-----TATACATACCATTTATGTAAGAGAGACAGCAATCGCTGTACG 852
DB 134 AAAAAGAGTCCAAATTTATATCTTCACTTAATGTCAGAGAGATGCTCAACTACTATATG 193

QY 853 GCTTTTCACAAGGAGAGAAAAATGCTGTTTACGMAAATCCTGAATGTTACGGGGATCG 912
DB 194 GCTTTTAAATGAAGAGGAAAAAGAAATGTTCTTAGCTTAATAAAGTGACTGGGATAG 253

QY 913 GCCCAAAAGGAGCGCTTTCGATCCTCGCTTCCGGCGATCCGGGAGCGGTGATTTGAAGCGA 972
DB 254 GACCTAAATCAGCTTTAGCGATACCTTCTTCAAGTACACACACATGAAGTTAAATTTGGCTA 313

QY 973 TCGAGATGAGGAGCAAGCATTTCTCGTCAAAATTTCCGGCGGTAGGCAAAAAACGGCAA 1032
DB 314 TCGAAATGAAATGATGCTTATTTAACACAATTTCTTGGGATAGGAAAGAAACTGCAA 373

QY 1033 GGCGAGATCATCTTGACCTGAAA 1055
DB 374 GACAAATGCTGTAGATTAAAA 396

RESULT 15
ABZ42153
ID ABZ42153 standard; DNA; 609 BP.
XX
AC ABZ42153;
XX
XX
DT 04-MAR-2003 (first entry)
XX
DE Streptococcus pneumoniae polynucleotide SEQ ID NO 7.
XX
KW Streptococcus pneumoniae; infection; otitis media; antibacterial;
XX diagnosis; gene therapy; gene; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200283855-A2.
XX
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-US011524.
XX
PR 16-APR-2001; 2001US-0283948P.
XX 18-APR-2001; 2001US-0284443P.
XX
PA (AMCY ) AMERICAN CYANAMID CO.
XX
PI Zagursky RJ, Masi AW, Green BA, Chakravarti DN, Russell DP;
PI Wooters JL;
XX
DR WPI; 2003-093010/08.
XX P-PSDB; ABP81305.
XX
PT New Streptococcus pneumoniae polynucleotides, useful for treating or
PT preventing S. pneumoniae infections, or non-systemic diseases, e.g.
XX otitis media, which are induced or exacerbated by S. pneumoniae.
XX
XX Claim 1; Page 183; 1091pp; English.
XX
```

The invention relates to isolated polynucleotides (ABZ72147-ABZ42522) of a *Streptococcus pneumoniae* genomic sequence, a fragment or degenerate variant of the polynucleotide or a nucleic acid sequence 95% identical to one of the polynucleotides. The *S. pneumoniae* polynucleotides and encoded polypeptides (ABP81299-ABP81674) are useful for treating or preventing *S. pneumoniae* infections or non-systemic diseases, e.g. otitis media, which are induced or exacerbated by *S. pneumoniae*. These are also useful for detecting *S. pneumoniae* in a biological sample or diagnosing *S. pneumoniae* infection in a subject. The polynucleotides have antibacterial activity and are useful in gene therapy

Sequence 609 BP; 184 A; 119 C; 146 G; 160 T; 0 U; 0 Other;

Query Match 8.5%; Score 89.8; DB 8; Length 609;
Best Local Similarity 54.0%; Pred. No. 1.3e-16;
Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY 675 ACGTTGATCGAATTCGTAAAGGGACGATTGATTATGTCGCCCAATATATTTGTCATT 734
DB 13 ACTATGTACGCATATTTAAAGGAATCATTACCAAAATTTACTGCCAAATACATTTGTTCTT 72

QY 735 GAAACCGCGGGATCGGCTAT-----CAGATCTTTCACGCCAAATCCGTTTATTATTAAG 788
DB 73 GAAACCAATGGTATTGGTTATATCTCTGATGTGGCAATCCTTATGCTTATTCAGGTGAG 132

QY 789 AANAACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGAGACACGAATCGGCTG 848
DB 133 GTTAATCAGGAGGCTCAGATTTATGTGCATCAGTTTGTGCGTAGGACGCCCATTTGCTT 192

QY 849 TACGGCTTTTCGACAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGG 908
DB 193 TATGGATTTGCTCAGAGGATGAGAAAAAGCTCTTCTTAGTCTGATTCGGTCTCTGGG 252

QY 909 ATCGGCCCAAAAGAGCGCTTTCGATCCTCGCTTCCGGCGATCCGGGAGCGGTGATTGAA 968
DB 253 ATTGGTCTGTATCAGCTCTTGTCTATTATGCTGCTGATGACAAATGCTGGCTTGGTTCAA 312

QY 969 GCGATCGAATGAGGACGAGCATTTCTCGTCAAAATTTCCCGCGGTAGGCAAAAAACG 1028
DB 313 GCCATTGAAACCAAGAACATCACCTACTTGACCAAGTTCCCTTAAATTTGGCAAGAAACA 372

QY 1029 GCAAGGCGAGATCATCTTGACCTGAAA 1055
DB 373 GCCCAGCAGATGCTGCTGGACTTGGAA 399

Search completed: February 27, 2006, 06:44:25
Job time : 751 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:32:27 ; Search time 1034 Seconds
(without alignments)
8437.325 Million cell updates/sec

Title: US-10-510-408-1
Perfect score: 1055
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacctgaaa 1055

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1055	100.0	1055	9	US-10-510-408-1
2	459	43.5	459	3	US-09-974-300-3090
3	195.4	18.5	546	3	US-09-974-300-1616
4	137	13.0	595	3	US-09-974-300-5972
5	122	11.6	495269	7	US-10-398-221-8
6	122	11.6	3011208	7	US-10-398-221-2058
7	108.2	10.3	1754382	9	US-10-501-282-6651
8	104.8	9.9	654	9	US-10-501-282-5561
9	104.8	9.9	654	9	US-10-501-282-5563
10	103.4	9.8	615	7	US-10-474-776-7
11	89.8	8.5	609	7	US-10-474-776-7
12	89.8	8.5	14273	2	US-08-961-527-40
13	89.8	8.5	14273	7	US-10-158-844-40
14	89.8	8.5	2162598	8	US-10-472-928-4979
15	89	8.4	591	8	US-10-472-928-167
16	89	8.4	7159	2	US-08-781-986A-302
17	89	8.4	7159	7	US-10-329-624-302
18	88.2	8.4	609	9	US-10-617-320-1114
19	87.4	8.3	603	8	US-10-857-625-351
20	87.4	8.3	11451	8	US-10-857-625-95
21	84.8	8.0	2121	7	US-10-398-221-3601
22	65.6	6.2	10996	3	US-09-070-927A-202
23	61.4	5.8	45613	7	US-10-672-787-22

ALIGNMENTS

RESULT 1

US-10-510-408-1									
; Sequence 1, Application US/10510408									
; Publication NO. US20050221423A1									
; GENERAL INFORMATION:									
; APPLICANT: Jorgensen, Steen Troels									
; APPLICANT: Olsen, Peter Bjarke									
; APPLICANT: Andersen, Jens Tonne									
; APPLICANT: Rasmussen, Michael Dolberg									
; TITLE OF INVENTION: Improved Bacillus Host Cell									
; FILE REFERENCE: 10295.204-US									
; CURRENT APPLICATION NUMBER: US/10/510,408									
; CURRENT FILING DATE: 2004-10-05									
; NUMBER OF SEQ ID NOS: 191									
; SOFTWARE: PatentIn version 3.3									
; SEQ ID NO 1									
; LENGTH: 1055									
; TYPE: DNA									
; ORGANISM: Bacillus licheniformis									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (34)..(552)									
US-10-510-408-1									
Query Match 100.0%; Score 1055; DB 9; Length 1055;									
Best Local Similarity 100.0%; Pred. No. 3.6e-304; Mismatches 0; Indels 0; Gaps 0;									
Matches 1055; Conservative 0;									
Qy	1	CCTTAAGGTAAGGCAAAAAAGAGGTGATTTGATGTACAGCGGCAAGTTCCTTCAAAATC	60						
Db	1	CCTTAAGGTAAGGCAAAAAAGAGGTGATTTGATGTACAGCGGCAAGTTCCTTCAAAATC	60						
Qy	61	GGTTTATTGCTTTATTTGGAAGTCTGTGGCCGCGCTTCACTTGAGGCTTTGGCC	120						
Db	61	GGTTTATTGCTTTATTTGGAAGTCTGTGGCCGCGCTTCACTTGAGGCTTTGGCC	120						
Qy	121	GAAGAAGCGGCTTAAAGTTCAATCCAGTTGAAAAGTTTATCTGGACGACGCTTGA	180						
Db	121	GAAGAAGCGGCTTAAAGTTCAATCCAGTTGAAAAGTTTATCTGGACGACGCTTGA	180						
Qy	181	ATTGAGATAAAGTACAGGCGCTCGCACACTGGAAGACTTTAAAGCTCTTTATAAGGG	240						
Db	181	ATTGAGATAAAGTACAGGCGCTCGCACACTGGAAGACTTTAAAGCTCTTTATAAGGG	240						
Qy	241	TGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTCGCAAAAGCTGGACGACATT	300						
Db	241	TGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTCGCAAAAGCTGGACGACATT	300						

Db 241 TGGCAGCTCATCGATCAGAAAAAGGGTTTATTCTGTTTCCGAAACAGGTGGACGACATT 300
Qy 301 TCTCCCTCAGCAAAACAAACGGTTTATATCGAGTGAAGTGAAGTGGCGTGAATTCGACT 360
Db 301 TCTCCCTCAGCAAAACAAACGGTTTATATCGAGTGAAGTGAAGTGGCGTGAATTCGACT 360
Qy 361 TTTACGGTGGCCGGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATA 420
Db 361 TTTACGGTGGCCGGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATA 420
Qy 421 AAGCGGTGGAAAGCGGATGCGGATGATCTGCGCAAAAGGATACCATACCGCACGAAA 480
Db 421 AAGCGGTGGAAAGCGGATGCGGATGATCTGCGCAAAAGGATACCATACCGCACGAAA 480
Qy 481 AAGGAATTGACATGTCTTGAAGCCGTAAATCATCCGAGAGCAACATCATGTAGAA 540
Db 481 AAGGAATTGACATGTCTTGAAGCCGTAAATCATCCGAGAGCAACATCATGTAGAA 540
Qy 541 GATATGAAGACATGACGCTGTTATGCTCTTTTTCAGCTGCAGACAGAAAGCTTTTTCAGG 600
Db 541 GATATGAAGACATGACGCTGTTATGCTCTTTTTCAGCTGCAGACAGAAAGCTTTTTCAGG 600
Qy 601 AACATATGTTAACTTTTTCATCTAGCTTTGCTGTTTGTGTTTACAATGAAGACAGTC 660
Db 601 AACATATGTTAACTTTTTCATCTAGCTTTGCTGTTTGTGTTTACAATGAAGACAGTC 660
Qy 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGAGACGATTTGATATGATCGCCCC 720
Db 661 AAAGAGGTGAATGAACGTTGATCGAATTCGTTAAAGGAGACGATTTGATATGATCGCCCC 720
Qy 721 ATATATATGTTCAATGAAACGCGGATCGGCTATCAGATCTTCAAGCCAAATCCGTTTA 780
Db 721 ATATATATGTTCAATGAAACGCGGATCGGCTATCAGATCTTCAAGCCAAATCCGTTTA 780
Qy 781 TTTATAAGAAAAACAGCAAGAAAAACAATCTATACATCATATGTAAGAGAACACGA 840
Db 781 TTTATAAGAAAAACAGCAAGAAAAACAATCTATACATCATATGTAAGAGAACACGA 840
Qy 841 ATGCCCTGTACGGCTTTTCGACAGGGAGAAAAAATGTTTACGAAAAATGCTGAATG 900
Db 841 ATGCCCTGTACGGCTTTTCGACAGGGAGAAAAAATGTTTACGAAAAATGCTGAATG 900
Qy 901 TTACGGGATCGGCCCAAAAGAGCGCTTCGATCTCGCTTCGGCGATCCGGAGCGG 960
Db 901 TTACGGGATCGGCCCAAAAGAGCGCTTCGATCTCGCTTCGGCGATCCGGAGCGG 960
Qy 961 TGATTGAAGCGATCAGAAATGAGGACGAATTTCTCGTCAAAATTTCCCGCGTAGGCA 1020
Db 961 TGATTGAAGCGATCAGAAATGAGGACGAATTTCTCGTCAAAATTTCCCGCGTAGGCA 1020
Qy 1021 AAAAAACGGCAAGGACGATCATCTTGTACCTGAAA 1055
Db 1021 AAAAAACGGCAAGGACGATCATCTTGTACCTGAAA 1055

RESULT 2

US-09-974-300-3090
; Sequence 3090, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3090
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-3090

Query Match 43.5%; Score 459; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 3.7e-126; Indels 0; Gaps 0;

Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 65 TATTGCTTATTGGAAGTCTGCTGGCCGCTCAGCTTTTTCAGCTGGAGCCCTTGGCCGAAA 124
Db 1 TATTGCTTATTGGAAGTCTGCTGGCCGCTCAGCTTTTTCAGCTGGAGCCCTTGGCCGAAA 60
Qy 125 AGCCGGCTAAAGTTCAAAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 184
Db 61 AGCCGGCTAAAGTTCAAAATCCAGCTTGAAGAGTTTATCTGGACGGAGACGTTGGAATTG 120
Qy 185 AGAATAAAGTAGAGCGCTCGCACATCGAAGACCTTTAAAGCTCTTTTAAAGGTTGC 244
Db 121 AGAATAAAGTAGAGCGCTCGCACATCGAAGACCTTTTAAAGCTCTTTTAAAGGTTGC 180
Qy 245 AGCTCATCGATCAGAAAAAGGGTTTATCTGTTTTCGCAAAACAGGTGACGACATTTCTC 304
Db 181 AGCTCATCGATCAGAAAAAGGGTTTATCTGTTTTCGCAAAACAGGTGACGACATTTCTC 240
Qy 305 CCCTCAGCAAAACAAACGGTTTATATCGAGTGAAGTGAAGTGGCGTGAATTCGACTTTTC 364
Db 241 CCCTCAGCAAAACAAACGGTTTATATCGAGTGAAGTGAAGTGGCGTGAATTCGACTTTTC 300
Qy 365 ACGGTCGCCCGCGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATAAAGC 424
Db 301 ACGGTCGCCCGCGGATCTTATCAGAACCCATTCAATCGTTTTCAGATTGATATAAAGC 360
Qy 425 GCCTGGAAGCGGATGCGGATGATCTGCGCAAAAGGATACCATACCGCACGAAAAAGG 484
Db 361 GCCTGGAAGCGGATGCGGATGATCTGCGCAAAAGGATACCATACCGCACGAAAAAGG 420
Qy 485 AATTGGAACATGTCATTTGAAGCCGTAAATCATCCGGAA 523
Db 421 AATTGGAACATGTCATTTGAAGCCGTAAATCATCCGGAA 459

RESULT 3

US-09-974-300-1616
; Sequence 1616, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1616
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(546)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-1616
Query Match 18.5%; Score 195.4; DB 3; Length 546;
Best Local Similarity 85.5%; Pred. No. 2.7e-47;

; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua
; US-10-398-221-2058

Query Match 11.6%; Score 122; DB 7; Length 3011208;
Best Local Similarity 59.4%; Pred. No. 3.2e-23;
Matches 228; Conservative 0; Mismatches 150; Indels 6; Gaps 1;

QY 678 TTGATCGAATTGTTAAAGGAGCAGATTGATTATGATCGCCCCCAATATATTGTCTAATTGAA 737
Db 1573493 TTGTACGATTACATAAAAGGAATAGTAAACGATTACACAGAAATATATTGTTGTCGAA 1573434

QY 738 AACGCGGGATCGGCTATCAGATCTTCAGCCCAATCCGTTTATTATAAGA-----AA 791
Db 1573433 ACAGACAAATCGGCTATCAATAATTACAGAAACCCATTTCCTTCAACGATTAGAA 1573374

QY 792 AACGCAAGAAACAATCTATACATACCAATATTGTAAGAGAGACACGAATGCGCTGTAC 851
Db 1573373 GGTACTGAAGCCAGTCTTTTGTATCAATGTTTCGGGAGACAAATTCCTTGTTT 1573314

QY 852 GCGTTTTCGACAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTTACGGGGATC 911
Db 1573313 GGTTCAGTCAACAGAGACGCTTATCTGTTCAAAAAATATTAAAGCGTTTCGGGTATT 1573254

QY 912 GCCCAAGAGGCGCTTCGATCTTCGCTCCGGCGATCCGGGAGCGGTGATTGAAGCG 971
Db 1573253 GGGCCAAAAGTGCCTAGCTATTATAGCTTCGGAGATGTAGTACCACTTATTACTGCA 1573194

QY 972 ATCGAGATGAGGAGCAGAGATTTCTCGTCAAAATTTCCGGGTAGGCAAAAAACGGCA 1031
Db 1573193 ATCGAATCAGAGATGATGTTTATTGTACTAAATTTCAAGTGTGCTGTAATAAACAGCA 1573134

QY 1032 AGGCAGATCATCTTGACCTGAAA 1055
Db 1573133 CGTCAAAATTATCTTGATTAAAA 1573110

RESULT 7
US-10-501-282-6651/c
; Sequence 6651, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6651

; LENGTH: 1754382
; TYPE: DNA
; ORGANISM: Alloiococcus otitidis
; US-10-501-282-6651

Query Match 10.3%; Score 108.2; DB 9; Length 1754382;
Best Local Similarity 56.3%; Pred. No. 3.1e-19;
Matches 247; Conservative 0; Mismatches 183; Indels 9; Gaps 2;

QY 623 CTAGCTTTGCTGTTTGTGTTTACAATGAAGAGCAGTCAAGAGGTGAATGAACCTTGAT 682
Db 1473455 CTAGGTTTATCTGTGTTATAATAAATGACTATTGGAAAAAGTGGTGAAGTA---TGTA 1473399

QY 683 CGAATTCGTAAGAGGACGATTGATTATGATCGCCCCCAATATATTGTCATTGAAAAACGG 742
Db 1473398 TGAATACATGTTGAGGACAGTCGTTGATGTCCAACCCAGCTACTTGGTTTTTACAAGTGGG 1473339

QY 743 CGGATCGGCTATCAGATCTTCAGCCAAATCCGTTTATTATTATAAGAAAAA-----CAG 796
Db 1473338 GGGAGTGGCTACCACTTTTAAATGGCTAACCCCTTCGCGCTGAACGACCGGCTGGCCA 1473279

QY 797 CAAGAAACAATCTATACATACCAATATTGTAAGAGAGACACGAATGCGCTGTACGGCTT 856
Db 1473278 AGAAGTGAAGATTTACCTTTACCTGAATGTTAGTCAAGACCAGCTTAGCCTCTTTGGTTT 1473219

QY 857 TTCGACAGGGAAGAAAAATGCTGTTTACGAAATGCTGAATGTTTACGGGATCGGCC 916
Db 1473218 TCCCGCCAGGATGAAAAAGCTTGTTCCTAAAATTAATCAATGTCCTCAGGTAATGGGGC 1473159

QY 917 AAAAGGAGCGCTTCGATCTTCGCTTCGCGGATCCGGGAGCGGTGATTGAAGCGATCGA 976
Db 1473158 CAAAAGTGTCTTGGCTATCTTGGCCAAATGAAGACCATGCTGGCTTAATCCAGGCGATTGA 1473099

QY 977 GAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGGCTAGGCAAAAAACGGCAAGCA 1036
Db 1473098 AAATGAAGACGACAGATTTTTCGCAACGCTTTTCCGGGGTAGGAAAAAACAAGCCTCTCA 1473039

QY 1037 GATCATCTTCACCTGAAA 1055
Db 1473038 GATTGTTCTGGACCTCAAA 1473020

RESULT 8
US-10-501-282-5561
; Sequence 5561, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5561
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Alloiococcus otitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (49) .. (651)
; US-10-501-282-5561

Query Match	9.9%;	Score 104.8;	DB 9;	Length 654;
Best Local Similarity	56.0%;	Pred. No. 3.6e-20;		
Matches 243;	Conservative 0;	Mismatches 182;	Indels 9;	Gaps 2;

Qy	628	TTTGCCCTGTTTGTGTTTACAATGAAGAGCAGTCAAAGAGGTCGAATGAACGTTGATCGAAT	687
Db			
Db	2	TTTATCTGTGGTATTAATAATGACTATTGGAAAAAGTGGTGAAGTA---TGTATGAAT	58
Qy	688	TCGTAATAAGGGACGATTGATTATGTATCGCCCCCAATATATTGTCAATGAAAAACGCGCGGA	747
Db			
Db	59	ACATGGTAGGGACAGTCGTGTGATGTCCAACCCAGCTACTTGGTTTTACAGTGGGGGAG	118
Qy	748	TCGGCTATCAGATCTTCACGCCAAATCGTTTATTTATTAAGAAAA-----ACACGAAG	801
Db			
Db	119	TGGGCTACACCTTTTAATGCGTAAACCCCTTCGCGCTGAAACGACCGGCTGGGCCAAGAAG	178
Qy	802	AAACAATCTATACATACCATTTATGTAAGAGAGACACGAATCGCGTGTACGGCTTTTCGA	861
Db			
Db	179	TGAAGATTACCTTTACCTGAATGTTAGTCAAGACCAGCTTAGCCCTCTTTGGTTTTCGCC	238
Qy	862	CAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCCAAAAG	921
Db			
Db	239	GCCAGATGAATAAGCTTGTTCTTAAATTAATCAATGCTCAGGATTTGGGCCCAAAA	298
Qy	922	GAGCGCTTTCGATCTCGCTCCCGCGATCCGGGAGCGGTGATTGAAGCGATCGAGAATG	981
Db			
Db	299	GTGCTTTGGCTATCTCGGCAATGAAGACCATGCTGGCTTAATCCAGGCGATTGAAAATG	358
Qy	982	AGGACGAAGCATTTCTCGTCAAAATTTCCCGGCTAGGCAAAAAACGGCAAGGCGATCA	1041
Db			
Db	359	AAGACCGACAGTTTTCGCAACGCTTTCCCGGGTAGSAAAAAAAACAGCCTTCTCAGATTG	418
Qy	1042	TCCTTGACCTGAAA	1055
Db			
Db	419	TTCTGGACCTCAA	432

RESULT 9
US-10-501-282-5563
; Sequence 5563, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5563
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Alloiococcus otitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22)..(651)
US-10-501-282-5563

Query Match	9.9%;	Score 104.8;	DB 9;	Length 654;
Best Local Similarity	56.0%;	Pred. No. 3.6e-20;		
Matches 243;	Conservative 0;	Mismatches 182;	Indels 9;	Gaps 2;

Qy	628	TTTGCCCTGTTTGTGTTTACAATGAAGAGCAGTCAAAGAGGTCGAATGAACGTTGATCGAAT	687
----	-----	--	-----

```
Db 134 AAAAAGAGTCCAAATTTATATCTTACCTAAATTTGTACGAGAGATGCTCAACTACTATATG 193
QY 853 GCTTTTCGACAGGGAAGAAATGCTGTTTACGAAATGCTGAATGTTACGGGGATCG 912
Db 194 GCTTTATTAATGAAGAGGAAAGAAATGTTTCTTAGCTTAATAAAGTGAAGTGGATAG 253
QY 913 GCCCAAAAGGAGCGCTTCGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAAGCGA 972
Db 254 GACCTAAATCAGCTTTAGCGATCTTCTTCAAGTACACCATGAAGTTAAATGGCTA 313
QY 973 TCGAGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGGTAGGCAAAAACGGCAA 1032
Db 314 TCGAAATGAAATGATGCTTATTAAACAAATTTCTCGGATAGAAAGAAACTGCAA 373
QY 1033 GCAGATCATCTTGAACCTGAAA 1055
Db 374 GACAAATGTTGTAGATTAAAA 396

RESULT 11
US-10-474-776-7
; Sequence 7, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLY
; FILE OF INVENTION: ANTIGENS AND USES THEREOF
; FILE REFERENCE: AM100649-PT
; CURRENT APPLICATION NUMBER: US/10/474,776
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-7

Query Match 8.5%; Score 89.8; DB 7; Length 609;
Best Local Similarity 54.0%; Pred. No. 1.1e-15;
Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY 675 ACGTTGATCGAATTCGTAAGGAGCGATTGATTATGATCGCCCAATATATTTGTCATT 734
Db 13 ACTATGTACGCATATTTAAAGGAATCATTAACAAATTTACTGCCAAATACATTTGTTCT 72
QY 735 GAAAACGGCGGATCGGCTAT-----CAGATCTTTCAGCCCAATCCCGTTTATTATAAG 788
Db 73 GAAACCAATGGTATTGGTTATATCTGTCATGTGGCAATCTTATGCGCTATTTCAGGTCAG 132
QY 789 AAAACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGAGACACGAATGCGGTG 848
Db 133 GTTAATCAGGAGGCTCAGATTTATGTGCATCAGGTTGTGCGTGAGGACGCCCATTTGCTT 192
QY 849 TACGGCTTTTCGACAGGGAAGAAATGCTGTTTACGAAATGCTGAATGCTACGGGG 908
Db 193 TATGGAATTCGCTCAGAGGATGAGAAAGAGCTTTTCTTAGTCTGATTTTCGCTCTCGGG 252
QY 909 ATCGGCCCCAAAGGAGCGCTTCGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAA 968
Db 253 ATTGGTCTGTATCAGCTCTTGTCTATTATCGTGTGATGACATGCTGGCTTGGTTCAA 312
QY 969 GCGATCGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGGTAGGCAAAAACG 1028
Db 313 GCCATTGAAACCAAGAACATCACCTACTTTGACCAAGTTCCCTAAAATTTGCCAAGAAACA 372
QY 1029 GCAAGGCAGATCATCTTGCACCTGAAA 1055
Db 373 GCCACGACAGATGGTGTGACTTGGAA 399

RESULT 12
US-08-961-527-40
```

```
; Sequence 40, Application US/08961527
; Publication No. US20020032323A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-40

Query Match 8.5%; Score 89.8; DB 2; Length 14273;
Best Local Similarity 54.0%; Pred. No. 6.4e-15;
Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY 675 ACGTTGATCGAATTCGTAAGGAGCGATTGATTATGATCGCCCAATATATTTGTCATT 734
Db 11348 ACTATGTACGCATATTTAAAGGAATCATTAACAAATTTACTGCCAAATACATTTGTTCT 11407
QY 735 GAAAACGGCGGATCGGCTAT-----CAGATCTTTCAGCCCAATCCCGTTTATTATAAG 788
Db 11408 GAAACCAATGGTATTGGTTATATCTGTCATGTGGCAATCTTATGCGCTATTTCAGGTCAG 11467
QY 789 AAAACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGAGACACGAATGCGGTG 848
Db 11468 GTTAATCAGGAGGCTCAGATTTATGTGCATCAGGTTGTGCGTGAGAGCGCCCATTTGCTT 11527
QY 849 TACGGCTTTTCGACAGGGAAGAAATGCTGTTTACGAAATGCTGAATGCTACGGGG 908
Db 11528 TATGGAATTCGCTCAGAGGATGAGAAAGAGCTTTTCTTAGTCTGATTTTCGCTCTCGGG 11587
QY 909 ATCGGCCCCAAAGGAGCGCTTCGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAA 968
Db 11588 ATTGGTCTGTATCAGCTCTTGTCTATTATCGTGTGATGACATGCTGGCTTGGTTCAA 11647
QY 969 GCGATCGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGGTAGGCAAAAACG 1028
Db 11648 GCCATTGAAACCAAGAACATCACCTACTTTGACCAAGTTCCCTAAAATTTGCCAAGAAACA 11707
QY 1029 GCAAGGCAGATCATCTTGCACCTGAAA 1055
Db 11708 GCCACGACAGATGGTGTGACTTGGAA 11734
```


RESULT 13

US-10-158-844-40
; Sequence 40, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
; APPLICANT: Kunech et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/158,844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB340P1D1
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-158-844-40

Query Match 8.5%; Score 89.8; DB 7; Length 14273;

Best Local Similarity 54.0%; Pred. No. 6.4e-15;

Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY	675	ACGTTGATCGAATTCGTAAAGGAGCGATTGATTGATCGCCCAATATATTGTCATT	734
Db	11348	ACTATGTACGCATATTTAAAGGAATCAATACCAAAATTAATGCCAAATACATTGTCCT	11407
QY	735	GAAGGCGGGATCGGCTAT-----CAGATCTTCACGCCAAATCCGTTTATTATAAG	788
Db	11408	GAACCAATGTTATGTTATATCTCTGATGTCGCAATCTTATGCTTATTCAGTCTAG	11467
QY	789	AAAAACAGCAAGAAACAATCTATACATCAATATGTAAGAGAACACGAAATCGCTG	848
Db	11468	GTTAATCAGAGGCTCAGATTATATGTCATCAGTTGTGCTGAGGACGCCATTTGCTT	11527
QY	849	TACGGCTTTTCGACAGGGAAGAAATGCTGTTTACGAAATGCTGAATGTTACGGGG	908
Db	11528	TATGGATTCGCTCAGAGGATGAGAAAAGCTCTTTCTAGTCTGATTTCCGGTCTCTGG	11587
QY	909	ATCGGCCAAAAGAGCGCTTCGATCCTCGCTTCGCGCATCCGGGAGCGGTGATTGAA	968
Db	11588	ATTGGTCTGTATCAGCTCTTGCTATATCGTCTGATGACAACTGCTGGTGGTTCAA	11647
QY	969	GCATCGAATGAGGACGAAGCATTTCTGTCAAATTTCCCGGCGTAGGCAAAAAACG	1028
Db	11648	GCCATTGAAACCAAGAACATCACCTACTTGACCAAGTTCCTAAATTTGCCAAGAAACA	11707
QY	1029	GCAAGCAGATCATCTTGCACCTGAA	1055
Db	11708	GCCACAGATGCTGCTGGACTTGGAA	11734

RESULT 14

US-10-472-928-4979
; Sequence 4979, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4979
; LENGTH: 2162598
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-472-928-4979

Query Match 8.5%; Score 89.8; DB 8; Length 2162598;

Best Local Similarity 54.0%; Pred. No. 1.1e-13;

Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY	675	ACGTTGATCGAATTCGTAAAGGAGCGATTGATTGATCGCCCAATATATTGTCATT	734
Db	65536	ACTATGTACGCATATTTAAAGGAATCAATACCAAAATTAATGCCAAATACATTGTCCT	65595
QY	735	GAAGGCGGGATCGGCTAT-----CAGATCTTCACGCCAAATCCGTTTATTATAAG	788
Db	65596	GAACCAATGTTATGTTATATCTCTGATGTCGCAATCTTATGCTTATTCAGTCTAG	65655
QY	789	AAAAACAGCAAGAAACAATCTATACATCAATATGTAAGAGAACACGAAATCGCTG	848
Db	65656	GTTAATCAGAGGCTCAGATTATATGTCATCAGTTGTGCTGAGGACGCCATTTGCTT	65715
QY	849	TACGGCTTTTCGACAGGGAAGAAATGCTGTTTACGAAATGCTGAATGTTACGGGG	908
Db	65716	TATGGATTCGCTCAGAGGATGAGAAAAGCTCTTTCTAGTCTGATTTCCGGTCTCTGG	65775
QY	909	ATCGGCCAAAAGAGCGCTTCGATCCTCGCTTCGCGCATCCGGGAGCGGTGATTGAA	968
Db	65776	ATTGGTCTGTATCAGCTCTTGCTATATCGTCTGATGACAACTGCTGGTGGTTCAA	65835
QY	969	GCATCGAATGAGGACGAAGCATTTCTGTCAAATTTCCCGGCGTAGGCAAAAAACG	1028
Db	65836	GCCATTGAAACCAAGAACATCACCTACTTGACCAAGTTCCTAAATTTGCCAAGAAACA	65895
QY	1029	GCAAGCAGATCATCTTGCACCTGAA	1055
Db	65896	GCCACAGATGCTGCTGGACTTGGAA	65922

RESULT 15

US-10-472-928-167
; Sequence 167, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 167
; LENGTH: 591

```

; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-472-928-167

Query Match      8.4%; Score 89; DB 8; Length 591;
Best Local Similarity 54.0%; Pred. No. 1.8e-15;
Matches 207; Conservative 0; Mismatches 170; Indels 6; Gaps 1;

QY 679 TGAATCGAATTCGTAAAGGAGGAGATTGATATGTCGCCCAATATATATGTCATTGAAA 738
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 TGTACGCATATTAAAGGAATCATACCAAAATTACTGCCAAATACATTGTTCTTGAAA 61
QY 739 ACGGGGGATCGCTAT-----CAGATCTTCACGCCCAATCCGTTTATTATTAAGAAA 792
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 CCAATGGTATGTTATATCCTGCAATGCGCAATCCTTATGCCATTACAGGTCAGGTTA 121
QY 793 ACAGCAAGAGAAACAATCTATACATACATTATGTAAGAGAAGACAGCAATGCGTACG 852
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 ATCAGGAGGCTCAGATTATGTCATCAGGTTGTGCGTGAGGACGCCCATTTGCTTTATG 181
QY 853 GCTTTTCGACAAGGGAGAGAAAATGCTGTTTACGAAAATGCTGAATGTACGGGATCG 912
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 GATTCGCTCAGAGGATGAGAAAAGCTCTTCTTAGTCTGATTCGGTCTCTGGGATG 241
QY 913 GCCCAAAAGGAGGCTTGGATCCTCGCTTCGGCGATCCGGGAGCGGTGATTGAAGCGA 972
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
242 GTCCGTGATCAGCTCTTGCTATATGCTGCTGATGACAAATGCTGGCTTGGTTCAAGCCA 301
QY 973 TCGAGAAATGAGGACGAAGCATTTCTCGTCAAAATTTCCCGCGTAGGCAAAAAACGGCAA 1032
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 TTGAACCAAGAACAATCACCCTACTTGACCAAGTTCCCTAAATTTGCAAGAAAACAGCCC 361
QY 1033 GGCAGATCATCCTTGACCTGAAA 1055
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 AGCAGATGGTGTGGACTTGGAA 384

```

Search completed: February 27, 2006, 08:12:58
Job time : 1047 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2006, 06:31:38 ; Search time 939 Seconds
(without alignments)
2395.790 Million cell updates/sec

Title: US-10-510-408-1
Perfect score: 1055
Sequence: 1 ccttaagtaaggcaaaaa.....agatcatccttgacgtgaaa 1055
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7209121 seqs, 1066183437 residues
Total number of hits satisfying chosen parameters: 14418242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_New.*
1: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
12: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1055	100.0	1522	8	US-10-510-386-123 Sequence 123, App
2	48.8	4.6	612	8	US-10-467-657-2695 Sequence 2695, Ap
3	48.8	4.6	612	8	US-10-467-657-2695 Sequence 6709, Ap
4	39	3.7	600	12	US-11-098-686-10086 Sequence 10086, A
5	39	3.7	1457619	12	US-11-098-686-8739 Sequence 8739, Ap
6	38.8	3.7	636	6	US-09-925-065A-102700 Sequence 102700,
7	38.8	3.7	636	6	US-09-925-065A-102701 Sequence 102701,
8	37.6	3.6	2436	8	US-10-793-626-4398 Sequence 4398, Ap
9	36.8	3.5	644	6	US-09-925-065A-295479 Sequence 295479,
10	36.8	3.5	611587	12	US-11-117-187-209 Sequence 209, App
11	36.6	3.5	364	6	US-09-925-065A-174583 Sequence 174583,
12	36.6	3.5	575	6	US-09-925-065A-147413 Sequence 147413,
13	36.4	3.5	580	6	US-09-925-065A-720480 Sequence 720480,
14	36.4	3.5	581	6	US-09-925-065A-477253 Sequence 477253,
15	36.4	3.5	2279	12	US-11-091-883-216 Sequence 216, App
16	36.4	3.5	212805	12	US-11-112-908-19 Sequence 19, Appl
17	36.2	3.4	1095	6	US-09-925-065A-42644 Sequence 42644, A
18	36.2	3.4	2092	8	US-10-750-185-64850 Sequence 64850, A
19	36.2	3.4	2092	8	US-10-750-623-64850 Sequence 64850, A
20	35.8	3.4	1911	8	US-10-750-185-32606 Sequence 32606, A

21	35.8	3.4	1911	8	US-10-750-623-32606 Sequence 32606, A
22	35.6	3.4	644	6	US-09-925-065A-295480 Sequence 295480,
23	35.4	3.4	385	6	US-09-925-065A-410936 Sequence 410936,
24	35.4	3.4	535	6	US-09-925-065A-503465 Sequence 503465,
25	35.2	3.3	588	6	US-09-925-065A-486701 Sequence 486701,
26	35.2	3.3	588	6	US-09-925-065A-486702 Sequence 486702,
27	35.2	3.3	588	6	US-09-925-065A-486703 Sequence 486703,
28	35.2	3.3	2571	6	US-09-925-065A-701231 Sequence 701231,
29	35.2	3.3	173115	12	US-11-112-908-65 Sequence 65, Appl
30	35.2	3.3	1082144	12	US-11-117-187-211 Sequence 211, App
31	35	3.3	525	6	US-09-925-065A-524053 Sequence 524053,
32	35	3.3	607	6	US-09-925-065A-724018 Sequence 724018,
33	35	3.3	631	6	US-09-925-065A-108779 Sequence 108779,
34	35	3.3	1095	6	US-09-925-065A-42642 Sequence 42642, A
35	35	3.3	1095	6	US-09-925-065A-42643 Sequence 42643, A
36	35	3.3	1603	6	US-09-925-065A-12675 Sequence 12675, A
37	35	3.3	394468	8	US-10-995-561-13473 Sequence 13473, A
38	34.8	3.3	525	6	US-09-925-065A-601441 Sequence 601441,
39	34.8	3.3	100000	12	US-11-124-367A-5026 Sequence 5026, Ap
40	34.6	3.3	558	6	US-09-925-065A-247578 Sequence 247578, A
41	34.6	3.3	581	6	US-09-925-065A-39072 Sequence 39072, A
42	34.6	3.3	584	6	US-09-925-065A-333708 Sequence 333708,
43	34.6	3.3	584	6	US-09-925-065A-333709 Sequence 333709,
44	34.6	3.3	48203	8	US-10-995-561-13378 Sequence 13378, A
45	34.4	3.3	615	6	US-09-925-065A-268787 Sequence 268787,

ALIGNMENTS

RESULT 1
US-10-510-386-123
; Sequence 123, Application US/10510386
; Publication No. US2005024922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 123
; LENGTH: 1522
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1019)
; US-10-510-386-123

Query Match 100.0%; Score 1055; DB 8; Length 1522;
Best Local Similarity 100.0%; Pred. No. 1.3e-276;
Matches 1055; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAAGGTAAGCAAAAAGGTAATTTGATGTACAGCGCAAGTTCAAAATC 60
Db 468 CCTTAAGGTAAGCAAAAAGGTAATTTGATGTACAGCGCAAGTTCAAAATC 527
QY 61 GGTATTGCTTTATTGGAAGTCTGCGCGCGCTCAGCTTCACTGGAGGCTTGGCC 120
Db 528 GGTATTGCTTTATTGGAAGTCTGCGCGCGCTCAGCTTCACTGGAGGCTTGGCC 587
QY 121 GAAAGCGCGCTTAAAGTTCAAATTCAGCTTTGAAAAGGTTTATCTGCACCGAGACGTTGGA 180
Db 588 GAAAGCGCGCTTAAAGTTCAAATTCAGCTTTGAAAAGGTTTATCTGCACCGAGACGTTGGA 647
QY 181 ATTGAGAATAAAGTAGAGCGCGCTCGCACACTGGGAAGACTTTAAAGCTCTTATAAGGG 240

Db 648 ATTGAGATAAAGTAGAGCCGCTCGCACACTGGAAGACTTTAAAGCTGCTTATAAAGG 707
Qy 241 TGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAACACAGGTGACGACATT 300
Db 708 TGGCAGCTCATCGATCAGAAAAAGGGTTTATCTGTTTCGCAACACAGGTGACGACATT 767
Qy 301 TCTCCCTCAGCAAAACAAAGGTTATATCGGAGTGACTGGAAGTGGCGTGAATTCGACT 360
Db 768 TCTCCCTCAGCAAAACAAAGGTTATATCGGAGTGACTGGAAGTGGCGTGAATTCGACT 827
Qy 361 TTTTCAGGTCGCGCGGATCTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 420
Db 828 TTTTCAGGTCGCGCGGATCTTATCAGAACCCATTCAATCGTTTTTTCAGATTGATATA 887
Qy 421 AAGCGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGACGAAA 480
Db 888 AAGCGCTGGAAGCGGATGCGGATGATCTGCGCAAGGATACCATACCGACGAAA 947
Qy 481 AAGGAATTTGAACATGTCAATGAAGCGGTAATCATCGGAAGCAACATCATGTAGAA 540
Db 948 AAGGAATTTGAACATGTCAATGAAGCGGTAATCATCGGAAGCAACATCATGTAGAA 1007
Qy 541 GATATGAAGACATGACGCTGTTATGCTTTTTTCAGCTGCAGACAGAAAGCTTTTTAGCG 600
Db 1008 GATATGAAGACATGACGCTGTTATGCTTTTTTCAGCTGCAGACAGAAAGCTTTTTAGCG 1067
Qy 601 AACATATGTTAACTTTTCAATCTAGCTTTGCTGTTTGTGTTTCAATGAAGACAGTC 660
Db 1068 AACATATGTTAACTTTTCAATCTAGCTTTGCTGTTTGTGTTTCAATGAAGACAGTC 1127
Qy 661 AAGAGGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATGATTATGTCGCCCC 720
Db 1128 AAGAGGTGAATGAACGTTGATCGAATTCGTAAGAGGACGATGATTATGTCGCCCC 1187
Qy 721 AATATATTGCTATTGAAAACGGCGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTA 780
Db 1188 AATATATTGCTATTGAAAACGGCGGATCGGCTATCAGATCTTCACGCCAAATCCGTTTA 1247
Qy 781 TTTATAAGAAAAACAGCAAGAAAAACAATCTATACATACCATTTATGTAAGAGAAGACAGA 840
Db 1248 TTTATAAGAAAAACAGCAAGAAAAACAATCTATACATACCATTTATGTAAGAGAAGACAGA 1307
Qy 841 ATGCGCTGTACGCTTTTCGCAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATG 900
Db 1308 ATGCGCTGTACGCTTTTCGCAAGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATG 1367
Qy 901 TTACGGGGATCGGCCCAAGAGGCGCTTCGATCTCGCTTCCGGCGATCCGGAGCGG 960
Db 1368 TTACGGGGATCGGCCCAAGAGGCGCTTCGATCTCGCTTCCGGCGATCCGGAGCGG 1427
Qy 961 TGATTGAACGGATCGGAATGAGGACGAAGCAATTTCTCGTCAAAATTTCCGGCGTAGGCA 1020
Db 1428 TGATTGAACGGATCGGAATGAGGACGAAGCAATTTCTCGTCAAAATTTCCGGCGTAGGCA 1487
Qy 1021 AAAAAACGCAAGGACGATCATCTTGCCTGAA 1055
Db 1488 AAAAAACGCAAGGACGATCATCTTGCCTGAA 1522

RESULT 2
US-10-467-657-2695
; Sequence 2695, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11

; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2695
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2695

Query Match 4.6%; Score 48.8; DB 8; Length 612;
Best Local Similarity 49.6%; Pred. No. 0.0032;
Matches 125; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
Qy 801 GAAACAATCTATACATACCATTTATGTAAGAGAAGACAGATGCGCTGTACGGCTTTTCG 860
Db 163 GTACAACATGTTTATCCAGCTTATCATTCGCGGAAGACGACATCTTTTATTTGTTTGGC 222
Qy 861 ACAAGGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCGGCCCAAAA 920
Db 223 ACTCGGAAGAACGCAAGACCTTCGCCCACTGATTAAGTTCGGCGGATCGCGCGGAAA 282
Qy 921 GGAGCGCTTGGGATCTCGCTTCCGGCGATCCGGAGCGGTGATTTGAAGCGATCGAGAAT 980
Db 283 ACGGCTTTGGGCAATTTTGTGCGCGATGACGCGACAGAGCTGGGCGGCGTTGCAGAA 342
Qy 981 GAGGACGAGCATTTTCTCGTCAAAATTTCCGGCGTAGGCAAAAAACCGCAAGGCAGATC 1040
Db 343 GAAGATGTCAAAACGCTCTCTCGCTCCGGGAATCGGCAAAAAAACCGCGAAGGTATG 402
Qy 1041 ATCCTTGACCTG 1052
Db 403 GTCCTGGAACATG 414

RESULT 3
US-10-467-657-6709
; Sequence 6709, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6709
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6709

Query Match 4.6%; Score 48.8; DB 8; Length 612;
Best Local Similarity 49.6%; Pred. No. 0.0032;
Matches 125; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
Qy 801 GAAACAATCTATACATACCATTTATGTAAGAGAAGACAGATGCGCTGTACGGCTTTTCG 860
Db 163 GTACAACATGTTTATCCAGCTTATCATTCGCGGAAGACGACATCTTTTATTTGTTTGGC 222
Qy 861 ACAAGGGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCGGCCCAAAA 920
Db 223 ACTCGGAAGAACGCAAGACCTTCGCCCACTGATTAAGTTCGGCGGATCGCGCGGAAA 282
Qy 921 GGAGCGCTTGGATCTCGCTTCCGGCGATCCGGAGCGGTGATTTGAAGCGATCGAGAAT 980

Db 283 ACGGCTTGGCATTGTCGGCGATGACGGCAGACGAGCTGGCGGGCGGTTCAGAA 342
QY 981 GAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAAAAGCGGACGAGATC 1040
Db 343 GAAGATGTAACACGCTCTCTCGCTCCGGGAATCGGCAAAAAAAGCGGCAACGTATG 402
QY 1041 ATCCTTGACCTG 1052
Db 403 GTCTTGAACCTG 414

RESULT 4
US-11-098-686-10086
; Sequence 10086, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10086
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10086

Query Match 3.7%; Score 39; DB 12; Length 600;
Best Local Similarity 46.5%; Pred. No. 1.5; Mismatches 0; Gaps 0;
Matches 126; Conservative 0; Indels 145; Indels 0; Gaps 0;
QY 785 TAAGAAAAACAGCAAGAAACAATCTATACATACCACTATGTAAGAGAAAGACACGAATGC 844
Db 120 TAATAAAGGAGAAACAAGTACGTTTTATATTTGTCTATATTTGACGGAGATATACAAGA 179
QY 845 GCTGTACGCTTTTCGACGAAGGAGAAATAATGCTGTTTACGAAATGCTGAATGTTAC 904
Db 180 ACTGTTTGGCTTTGAAACATGGATGAGCGACAAAGCTTTATTTGTTTAACTATCTATATC 239
QY 905 GGGGATCGGCCCAAAAGGAGCGCTTCGATCTCGCTTCGGCGATCCGGGAGCGGTAT 964
Db 240 TAAAGTCGGTGACGTACGACACTTAACATGCTTTCACTTTTCGACCAAAATGATCTAAG 299
QY 965 TGAAGCGATCGAAGATGAGGACGAAGCAATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAA 1024
Db 300 ACAATGTGTTATAGAAGAAGATATTTTCTTACTCTGCTGATCTGGAATTTGGAAGAA 359
QY 1025 AACGGCAGGCGAGATCATCTTGGACCTGAAA 1055
Db 360 GACAGCAACAATATTTTCTTGAACCTTAAA 390

RESULT 5
US-11-098-686-8739
; Sequence 8739, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395

; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8739
; LENGTH: 1457619
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8739

Query Match 3.7%; Score 39; DB 12; Length 1457619;
Best Local Similarity 46.5%; Pred. No. 35; Mismatches 145; Indels 0; Gaps 0;
Matches 126; Conservative 0; Indels 145; Indels 0; Gaps 0;
QY 785 TAAGAAAAACAGCAAGAAACAATCTATACATACCACTATGTAAGAGAAAGACACGAATGC 844
Db 1455272 TAATAAAGGAGAAACAAGTACGTTTTATATTTGTCTATATTTGACGGAGATATACAAGA 1455331
QY 845 GCTGTACGCTTTTCGACGAAGGAGAAATAATGCTGTTTACGAAATGCTGAATGTTAC 904
Db 1455332 ACTGTTTGGCTTTGAAACATGGATGAGCGACAAAGCTTTATTTGTTTAACTATCTATATC 1455391
QY 905 GGGGATCGGCCCAAAAGGAGCGCTTCGATCTCGCTTCGGCGATCCGGGAGCGGTAT 964
Db 1455392 TAAAGTCGGTGACGTACGACACTTAACATGCTTTCACTTTTCGACCAAAATGATCTAAG 1455451
QY 965 TGAAGCGATCGAAGATGAGGACGAAGCAATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAA 1024
Db 1455452 ACAATGTGTTATAGAAGAAGATATTTTCTTACTCTGCTGATCTGGAATTTGGAAGAA 1455511
QY 1025 AACGGCAGGCGAGATCATCTTGGACCTGAAA 1055
Db 1455512 GACAGCAACAATATTTTCTTGAACCTTAAA 1455542

RESULT 6
US-09-925-065A-102700/c
; Sequence 102700, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 10827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102700
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-102700

Query Match 3.7%; Score 38.8; DB 6; Length 636;
Best Local Similarity 50.5%; Pred. No. 1.7; Mismatches 92; Indels 0; Gaps 0;
Matches 94; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 489 TGAACATGCTATTGAAGCGCGTAAATCATCCGGAAGCCCAACATCATGTAGAGATATGAA 548
Db 208 TGCCCATGACTTTTGAAGAAGTTAAGTCAATCTGCAATATCAAGGTACTCAGCTTCTGCTGAA 149
QY 549 GACATGACGCTGTTATGCTTTTTCAGCTGCAGACAGAAGCTTTTTTAGCGAACAATATG 608

Db	148	TGCTTATCTCTTAAGTCTAATTGAGCAATATGAATAATTTTGGATGGAGATCTCATC	89
Qy	609	TTAACTTTTTCATCTAGCTTTGCCTGTTTGTGTACAAATGAAGAGCAGTCAAGAGGT	668
Db	88	TTAAATGTTCCCTCTACCTACTTTCTTAAATGTCCTATCAATGCAATGAACATTT	29
Qy	669	GAATGA	674
Db	28	CCTTGA	23
RESULT 7			
US-09-925-065A-102701/c			
; Sequence 102701, Application US/09925065A			
; Publication No. US20040181048A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, David G.			
; TITLE OF INVENTION: Identification and Mapping of Single			
; FILE REFERENCE: 108827.135			
; CURRENT APPLICATION NUMBER: US/09/925,065A			
; PRIOR FILING DATE: 2001-08-08			
; PRIOR APPLICATION NUMBER: US 60/243,096			
; PRIOR FILING DATE: 2000-10-24			
; PRIOR APPLICATION NUMBER: US 60/252,147			
; PRIOR FILING DATE: 2000-11-20			
; PRIOR APPLICATION NUMBER: US 60/250,092			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: US 60/261,766			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/289,846			
; PRIOR FILING DATE: 2001-05-09			
; NUMBER OF SEQ ID NOS: 957086			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 102701			
; LENGTH: 636			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-065A-102701			
Query Match			
Best Local Similarity 3.7%; Score 38.8; DB 6; Length 636;			
Matches 94; Conservative 0; Mismatches 92; Indels 0; Gaps 0;			
Qy	489	TGAACATGTCATGAAGCCGTAAATCATCCGAAACCAACATCATGTAGAGATATGAA	548
Db	208	TGCCCATGACTTTTGAAGAAGTTAAGTCAATTCGCAATAATCAAGGTACTCAGCTTGCTGAA	149
Qy	549	GACATGACGCTGTATGTCTTTTTCAGCTGCAGACAGAGCTTTTATAGCGAACATATG	608
Db	148	TGCTTATCTCTTCTAAGTCTAATTTGAGCAATATGAATAATTTTGGATGGAGATCTCATC	89
Qy	609	TTAACTTTTTCATCTAGCTTTGCCTGTTTGTGTACAAATGAAGAGCAGTCAAGAGGT	668
Db	88	TTAAATGTTTCCCTCTACCTACTTTCTTAAATGTCCTATCAATGCAATGAACATTT	29
Qy	669	GAATGA	674
Db	28	CCTTGA	23
RESULT 8			
US-10-793-626-4398			
; Sequence 4398, Application US/10793626			
; Publication No. US20050255478A1			
; GENERAL INFORMATION:			
; APPLICANT: KIMMERLY, WILLIAM JOHN			
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS			
; FILE REFERENCE: PU3480US			
; CURRENT APPLICATION NUMBER: US/10/793,626			
; PRIOR FILING DATE: 2004-03-04			
; PRIOR APPLICATION NUMBER: 60/164,258			
; PRIOR FILING DATE: 1999-11-09			
; NUMBER OF SEQ ID NOS: 4472			
; SOFTWARE: PatentIn Ver. 2.1			
; LENGTH: 2436			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence: synthetic			
; OTHER INFORMATION: nucleic acid sequence			
US-10-793-626-4398			
Query Match			
Best Local Similarity 3.6%; Score 37.6; DB 8; Length 2436;			
Matches 55; Conservative 0; Mismatches 29; Indels 0; Gaps 0;			
Qy	972	ATCGAAGATGAGGACGAGCATTTCTCGTCAAAATTTCCCGGCTAGGCAAAAAACGGCA	1031
Db	2	ATCGAAGATGAAATGATGCTTATTACACAATTTCTCTGGATAGGAAAGAAACTGCA	61
Qy	1032	AGGCAGATCATCCTTGACCTGAAA	1055
Db	62	AGACAAATGTGTAGATTAAAA	85
RESULT 9			
US-09-925-065A-295479/c			
; Sequence 295479, Application US/09925065A			
; Publication No. US20040181048A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, David G.			
; TITLE OF INVENTION: Identification and Mapping of Single			
; FILE REFERENCE: 108827.135			
; CURRENT APPLICATION NUMBER: US/09/925,065A			
; PRIOR FILING DATE: 2001-08-08			
; PRIOR APPLICATION NUMBER: US 60/243,096			
; PRIOR FILING DATE: 2000-10-24			
; PRIOR APPLICATION NUMBER: US 60/252,147			
; PRIOR FILING DATE: 2000-11-20			
; PRIOR APPLICATION NUMBER: US 60/250,092			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: US 60/261,766			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/289,846			
; PRIOR FILING DATE: 2001-05-09			
; NUMBER OF SEQ ID NOS: 957086			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 295479			
; LENGTH: 644			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-065A-295479			
Query Match			
Best Local Similarity 3.5%; Score 36.8; DB 6; Length 644;			
Matches 101; Conservative 1; Mismatches 108; Indels 0; Gaps 0;			
Qy	527	AACATCATGTAGAAGATATGAACACATGACGCTGTTATGTCTTTTTCAGCTGCAGACAG	586
Db	519	AAGAGCTTTCCAGAAATGGAACAATGCCACATTTTCTAATTTTGTGTTTGGAAAAATA	460
Qy	587	AAGCTTTTTAGCGAACATATGTTAACTTTTTCATCTAGCTTTGCCTGTTTGTGTAC	646
Db	459	AACTTTTGGTAAAAATATGTTATTATTAATTAACATGTAATGTTTATGATGTTATTTT	400
Qy	647	AATGAGAGCAGTCAAGAGGTGAATGAACGTTGATCGAATTCGTAAGGAGGACGATTGA	706
Db	399	AAATAAATTAATAAACAACATGTTTGAATTTGTTTAATTTCTAATATGTTAAATTTT	340
Qy	707	TTATGTATCGCCCAATATATTTGTCATTGA	736
Db	339	TTATTATKATCTTTTATTTTGTGAGACTGA	310

```
RESULT 10
US-11-117-187-209/c
; Sequence 209, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 209
; LENGTH: 611587
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-209

Query Match      3.5%; Score 36.8; DB 12; Length 611587;
Best Local Similarity 51.9%; Pred. No. 96;
Matches 83; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 139 CAAATCCAGCTGAAAGGTTTATCTGGACGGAGAGCTTGGAAATGAGATAAGTAGAG 198
Db 89346 CAAATGGATGTTAAATAATCGGTTCTTCAAGGAGAACTCGAAGAAAAAGTGACATGAAG 89287

QY 199 GCCGCTCGCACACTGGAAGCTTTAAAGCTGCTTATAAAGGCTGAGCTCATCGATCAG 258
Db 89286 CCACCACCGGCTGAGACACATTAATGCTCCCAATAAGTCTTAAAGCTTAAGAAAGCC 89227

QY 259 AAAAAGGGGTTTATCTGTTGTCGCAACACGGTGGAGACA 298
Db 89226 ATCTACGGTTTAAAGCAATCTCCGAGAGCGTGGTACCACA 89187

RESULT 11
US-09-925-065A-174583/c
; Sequence 174583, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174583
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-174583

Query Match      3.5%; Score 36.6; DB 6; Length 364;
Best Local Similarity 55.0%; Pred. No. 5.4;
Matches 93; Conservative 0; Mismatches 74; Indels 2; Gaps 1;
```

```
QY 477 GAAAAAGGAATTTGAACATGTCTAATGAAGCCGTAATAATCATCCGGAAGCCAAACATCATGT 536
Db 358 GGAATAAACATTTATAAGAAAGTCAGCTATGGTTAGAAAGTCAGCTTAATGAAGAATGATAT 299

QY 537 AGAAGATATGAAGACATGACGCTGTATGTCTTTTTCAGCTGCAG--ACAGAAGCTTTT 594
Db 298 TCAGGATATATATATATCTTTCTGTGTTTTTTTTTCTGTGTCTCTGTGCTGCTGTTT 239

QY 595 TTAGCGAACATATGTTAACTTTTTCATCTAGCTTTGCTGCTGTTTGTGT 643
Db 238 TTAATCACTACACACACCACCTTTTAATATATATATTTGATCCTTTTGT 190

RESULT 12
US-09-925-065A-147413/c
; Sequence 147413, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147413
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-147413

Query Match      3.5%; Score 36.6; DB 6; Length 575;
Best Local Similarity 53.1%; Pred. No. 6.6;
Matches 78; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 572 TTCAGCTCGACAGACAGAGCTTTTTCAGCAACATATGTTAACTTTTCACTTAGCTTTG 631
Db 481 TTCACCTTTCTGCAGTACCTGAAATAGCATACATTTTTCAGCTTTGCCATTCATAGTAC 422

QY 632 CCGTGTGTTGTGTACAATGAAGAGCAGTCAAGAGGTCGAAGAGGTCGAACGTTGATCGAATTCGT 691
Db 421 TCTGATGCTTTACATTTCTTCAGCTTTTAAATGAATTTTGTCTTTAACTAATAAGA 362

QY 692 AAAAGGGACGATTTGATTTATGTATCGCC 718
Db 361 AAAAGGCCCTTTTATATATACATCGCC 335

RESULT 13
US-09-925-065A-720480/c
; Sequence 720480, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
```


Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	103.4	9.8	615	3	US-09-134-001C-780	Sequence 780, App
2	89.8	8.5	14273	3	US-08-961-527-40	Sequence 40, Appl
3	89	8.4	7159	3	US-08-956-171E-302	Sequence 302, App
4	89	8.4	7159	3	US-08-781-986A-302	Sequence 302, App
5	88.2	8.4	609	3	US-09-107-433-1114	Sequence 1114, App
6	87.4	8.3	594	3	US-09-583-110-2142	Sequence 2142, Ap
7	74.4	7.1	696	3	US-09-134-000C-3189	Sequence 3189, Ap
8	65.6	6.2	633	3	US-09-489-039A-6958	Sequence 6958, Ap
9	61.4	5.8	624	3	US-09-540-236-555	Sequence 555, App
10	61.4	5.8	45613	3	US-09-596-002-22	Sequence 22, Appl
11	60.2	5.7	609	3	US-09-328-352-2228	Sequence 2228, Ap
12	60.2	5.7	657	3	US-09-543-681A-361	Sequence 361, App
13	50.6	4.8	3250	3	US-09-221-017B-960	Sequence 960, App
14	49	4.6	1830121	3	US-09-557-884-1	Sequence 1, Appl
15	49	4.6	1830121	3	US-09-643-990A-1	Sequence 1, Appl
16	49	4.6	1830121	3	US-10-158-865-1	Sequence 1, Appl
17	47.8	4.5	510	3	US-09-252-991A-13007	Sequence 13007, A
18	47.8	4.5	873	3	US-09-252-991A-12464	Sequence 12464, A
19	47.8	4.5	7218	2	US-08-233-463-14	Sequence 14, Appl
20	46	4.4	609	3	US-09-902-540-2478	Sequence 2478, Ap
21	46	4.4	15209	3	US-09-902-540-1110	Sequence 1110, Ap
22	43.4	4.1	558	3	US-09-583-110-2141	Sequence 2141, Ap
23	40.6	3.8	79350	3	US-09-949-016-12467	Sequence 12467, A
24	40.6	3.8	79351	3	US-09-949-016-12467	Sequence 12467, A
25	40.6	3.8	79351	3	US-09-949-016-16275	Sequence 16275, A

Db 314 TCGAAATGAAATGATCTTATTAAACAAATTCCTGGGTAGGAAGAAACTGCA 373
QY 1033 GCGAGATCATCTTGACCTGAAA 1055
Db 374 GACAAATTGTGTAGATTAAAA 396

RESULT 2
US-08-961-527-40
; Sequence 40, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-40

Query Match 8.5%; Score 89.8; DB 3; Length 14273;
Best Local Similarity 54.0%; Pred. No. 7.6e-18;
Matches 209; Conservative 0; Mismatches 172; Indels 6; Gaps 1;

QY 675 ACGTTGATCGAATTCGTAAGGAGGAGATTGATTGATGCCCAATATATTCGTCATT 734
Db 11348 ACTATGTAGCATATTTAAAGGATCATACCAAAATCTGCCAAATACATGTTCTT 11407

QY 735 GAAACGGCGGGATCGGCTAT-----CAGATCTTCACGCCAAATCCGTTTATTATAAG 788
Db 11408 GAAACCAATGGTATTGGTTATATCTGCAATGTGGCAATCTTTATGCTTATTCAGGTCAG 11467

QY 789 AAAACAGCAAGAACAATCTATACATACCATATGTAAGAGAGACACGATCGGCTG 848
Db 11468 GTTAATCAGGAGGCTCAGATTATGTGTCATCAGGTTGTGCGTGAGGACGCCAATTGCTT 11527

QY 849 TACGGCTTTTCACAGAGGAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTTACGGGG 908
Db 11528 TATGATTTTCGTCAGAGGATGAGAAAAGCTTTCTTAGTCTGATTCGCTCTCTGGG 11587

QY 909 ATCGGCCCCAAAAGGACGCTTCGATCTCGCTTCGGCGATCCGGGAGCGGTGATTGAA 968
Db 11588 ATTGGTCTGTATCAGCTCTTCTGCTATTATCGCTGCTGATGACAAATGCTGGCTTGGTTCAA 11647

QY 969 GCGATCGAGATGAGGACGAGCATTTCCTCGTAAATTTCCCGCGTAGGCAAAAAACG 1028
Db 11648 GCCATTGAACCAAGAAATCACCCTACTTGACCAAGTTCCTTAAATTTGGCAAGAAACA 11707

QY 1029 GCAAGGCGAGATCATCTTGGACCTGAAA 1055
Db 11708 GCCCAGCAGATGGTCTGGACTTGAA 11734

RESULT 3
US-08-956-171E-302
; Sequence 302, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 302:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7159 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 302:
US-08-956-171E-302

Query Match 8.4%; Score 89; DB 3; Length 7159;
Best Local Similarity 54.0%; Pred. No. 9.1e-18;
Matches 207; Conservative 0; Mismatches 170; Indels 6; Gaps 1;

QY 679 TGATCGAATTCGTAAGGAGCAGATTGATTGATCGCCCAATATATTTGTCATTGAAA 738
Db 3226 TGTACGCGTGTGTCAAAGGTAAGTTAACACATTATATATCTACACACGTTGTTGAAA 3285

QY 739 ACGCGGGATCGGCTATCAGATCTTCAGCCAAATCCGTTTATTATAAGAA-----AA 792
Db 3286 CTGCTGGTGTGGTTATGAAATTCAAACACCAATCTTATCGTTTTCAAAAGCATCTAG 3345

QY 793 ACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGAGACACGAATGCGCTGACG 852

Db 3346 ATCATGAAGTTTAAATTCATACATCTTTAAATGTTTCGTAAGATGCAAAATTTATGTTATG 3405
Qy 853 GCTTTTCGACAAAGGAGGAGAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGGATCG 912
Db 3406 GATTTTAGTAGTGAAGAGAGAAATATGTTCTTGAGTTTAAATTAAGTTACTGTTATG 3465
Qy 913 GCCAAAAGGAGCGCTTGGGATCCTCGCTTCCGGGATCCGGGAGCGGTGATTTGAAGCGA 972
Db 3466 GTCCGAAATCAGCTTTAGCTATTATTTAGGCAAGTACGCTTAATGAAGTAAACCGTGCA 3525
Qy 973 TCGAAGTGAAGACGAGCATTTCTCGTCAAAATTTCCGGGCTAGGCAAAAACGGCAA 1032
Db 3526 TTGAAATGAAGATGATGATTTAACTAAATTTCCAGGAATTTGTAAGAAAACGGCAA 3585
Qy 1033 GGCAGATCATCTTGACCTGAAA 1055
Db 3586 GACAGATTGCTCTTAGATTTAAAA 3608

RESULT 4

US-08-781-986A-302
; Sequence 302, Application US/08781986A
; Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 302:

SEQUENCE CHARACTERISTICS:

LENGTH: 7159 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-302

Query Match 8.4%; Score 89; DB 3; Length 7159;

Best Local Similarity 54.0%; Pred. No. 9,1e-18;

Matches 207; Conservative 0; Mismatches 170; Indels 6; Gaps 1;

Qy 679 TGATCGAATTCGTAAGGAGGAGTTGATGATGATCGCCCAATATATGTCATGTGAAA 738

Db 3226 TGTACGCGTATGCAAGGTAAAGTAAACACATTTATATCTACACACGTAGTTGTTGAAA 3285

Qy 739 ACGGGGATCGGCTATCAGATCTTCAGCCAAATCCGTTTATTATAGAA-----AA 792

Db 3286 CTGCTGGTGTGGTTATGAAATTTCAACACCAAAATTTCTTATCGTTTTCAAAAGCATCTAG 3345

Qy 793 ACAGCAAGAAACAATCTATACATACCATTTATGTAAGAGAAAGACACGAATGCGCTGTACG 852
Db 3346 ATCATGAAGTTTAAATTCATACATCTTTAAATGTTTCGTAAGATGCAAAATTTATGTTATG 3405
Qy 853 GCTTTTCGACAAAGGAGGAGAAAATGCTGTTTACGAAAATGCTGAATGTTACGGGGATCG 912
Db 3406 GATTTTAGTAGTGAAGAGAGAAATATGTTCTTGAGTTTAAATTAAGTTACTGTTATG 3465
Qy 913 GCCAAAAGGAGCGCTTGGGATCCTCGCTTCCGGGATCCGGGAGCGGTGATTTGAAGCGA 972
Db 3466 GTCCGAAATCAGCTTTAGCTATTATTTAGGCAAGTACGCTTAATGAAGTAAACCGTGCA 3525
Qy 973 TCGAAGTGAAGACGAGCATTTCTCGTCAAAATTTCCGGGCTAGGCAAAAACGGCAA 1032
Db 3526 TTGAAATGAAGATGATGATTTAACTAAATTTCCAGGAATTTGTAAGAAAACGGCAA 3585
Qy 1033 GGCAGATCATCTTGACCTGAAA 1055
Db 3586 GACAGATTGCTCTTAGATTTAAAA 3608

RESULT 5

US-09-107-433-1114

; Sequence 1114, Application US/09107433

; Patent No. 6800744

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

THERAPEUTICS

NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>

SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/ 085131

FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1114:

SEQUENCE CHARACTERISTICS:

LENGTH: 609 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1....609

FOR DIAGNO

Db	209	ATGATCGTGATATTTTTAGAAAGCTGATCAAAATTAACGCGGTAGGTGCTTAAAAATGCGGT	268
Qy	928	TTGCGATCCCTCGCTTCGCGGATCCGSGAGCGGTGATTGAAGCGATCGAATGAGGACG	987
Db	269	TAGCAATGCTTTCAACCTTAAGCTTGCCGAGATTAAACACGCTATTGAAACACAGATTATG	328
Qy	988	AAGCATTTCTCGTCAAAATTTCCGCGGTAGGCAAAAAACGGCAAGCGAGATCATCTTGG	1047
Db	329	ATGCAGCTTTGGTGGCGTGTTACCGGGCATTGGCAAAAAAACTGCTCAGCGTATTTTGGCTTG	388
Qy	1048	ACCTGAA 1054	
Db	389	ATTATAA 395	
RESULT 10			
US-09-596-002-22			
; Sequence 22, Application US/09596002			
; Patent No. 6632636			
; GENERAL INFORMATION:			
; APPLICANT: Lagace, Robert, E.			
; APPLICANT: Patterson, Chandra			
; APPLICANT: Berg Kim, L.			
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME			
; FILE REFERENCE: PM-0008-4 US			
; CURRENT APPLICATION NUMBER: US/09/596,002			
; CURRENT FILING DATE: 2000-06-16			
; PRIOR APPLICATION NUMBER: 60/140,121			
; PRIOR FILING DATE: 1999-06-18			
; NUMBER OF SEQ ID NOS: 41			
; SOFTWARE: PERL Program			
; SEQ ID NO 22			
; LENGTH: 45613			
; TYPE: DNA			
; ORGANISM: M. catarrhalis			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Incyte template ID No. 6632636 22			
; PUBLICATION INFORMATION:			
US-09-596-002-22			
Query Match 5.8%; Score 61.4; DB 3; Length 45613;			
Best Local Similarity 53.0%; Pred.No.3.2e-08;			
Matches 131; Conservative 0; Mismatches 116; Indels 0; Gaps 0;			
Qy	808	TCATATACATACCATTTATGTAAGAGAGACACGAATCGCTGTACGGCTTTTCGACAAGGG	867
Db	9213	TCATATGTACATCATCATGTCGCGAGATGCCATAAATTTGTTTGGTTTTATGGATCGTC	9272
Qy	868	AGAAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATTCGGCCCAAAAGGAGCGC	927
Db	9273	ATGATCGTGATATTTTTTGAAGAAGCTGATCAAAATTAACGCGGTAGGTGCTTAAAAATGCGCT	9332
Qy	928	TTGCGATCTCGCTTCGCGGATCCGSGAGCGGTGATTGAAGCGATCGAGATGAGGACG	987
Db	9333	TAGCAATGCTTTCAACCTTAAGCTTGCCGAGATTAAACACGCTATTGAAACACAGATTATG	9392
Qy	988	AAGCATTTCTCGTCAAAATTTCCGCGGTAGGCAAAAAACGGCAAGCGAGATCATCTTGG	1047
Db	9393	ATGCAGCTTTGGTGGCGTGTTACCGGGCATTGGCAAAAAAACTGCTCAGCGTATTTTGGCTTG	9452
Qy	1048	ACCTGAA 1054	
Db	9453	ATTATAA 9459	
RESULT 11			
US-09-328-352-2228			
; Sequence 2228, Application US/09328352			
; Patent No. 6562958			
; GENERAL INFORMATION:			
; APPLICANT: Gary L. Breton et al.			

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2228
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-228

Query Match 5.7%; Score 60.2; DB 3; Length 609;
Best Local Similarity 51.7%; Pred. No. 5.8e-09;
Matches 137; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 787 AGAAAAACAGCAAGAAACAATCTATACATACCATATATGTAGAGAGACACGAATGGC 846
Db |||||
Qy 125 AAAAAAGCCAGAAAGTTAGCTATGACGCGATTTGGTGGTGAAGATGCTCAGCAGC 184
Db |||||
Qy 847 TGTACGGCTTTTCGACAGGGAAGAAAAATGCTGTTTACGAAAATGCTGAATGTTACGG 906
Db |||||
Qy 185 TTTATGGTTTATGATGCTCAAGAGAAAACTATTTCCGCACCTTACTAAAAAGTAAATG 244
Db |||||
Qy 907 GGATCGGCCCAAGAGGCGCTTGCATCCTCGCTTCGGCGATCCGGGAGCGGTGATTG 966
Db |||||
Qy 245 GTGGGGCCCAAAATGCGCATAGTATTTCTTACGTTAAGTGTGAATGCTGGTAC 304
Db |||||
Qy 967 AAGCGATCGAATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGCGTAGGCAAAAAA 1026
Db |||||
Qy 305 ACACAAATGAACATGACGATGTAATACTTTGGTTAAAGTTCCGGGTGTAGGCAAAAAA 364
Db |||||
Qy 1027 CGGCAAGCGAGATCATCTTGGACT 1051
Db |||||
Qy 365 CAGCTGAACGTTTAAATGATGAACT 389
Db |||||

RESULT 12

US-09-543-681A-361
; Sequence 361, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 361
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-361

Query Match 5.7%; Score 60.2; DB 3; Length 657;
Best Local Similarity 52.6%; Pred. No. 5.8e-09;
Matches 131; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 807 ATCTATACATACCATATGTAGAGAGACACGAGTGGCTGTACGGCTTTTCGACAGG 866
Db |||||
Qy 175 ATCTACACACAAATTTATTGTCGTGAAGATGCGCAATTACTATATGTTTAAATCAAAAA 234
Db |||||
Qy 867 GAAGAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGATCGGCCCAAGAGCG 926
Db |||||
Qy 235 CAAGAACGGGCATGTTTTCGTGAATTTGATTAAGTCAATGTTGTTAGGCGCTAAGCTAGCA 294
Db |||||
Qy 927 CTTGCGATCCTGCTTCCGGGATCCGGAGCGGTGATTTGAAGCGATCGAAGATGAGGAC 986
Db |||||
Qy 295 CTGGCTATTTTATCAGGAATGCTCGACGCCAAATTTGTACCGCAATTTGTCACCGCAATGAAAAATGAATCC 354
Db |||||

Qy 987 GAAGCATTTTCGTCAAATTTCCGGCGTAGGCAAAAAAAGCGCAAGSCAGATCATCTT 1046
Db |||||
Qy 355 ATCTCTCATTTAGTTAAATTAACCGGAGTGGTAAAAAACCAGGACGATTAGTGTT 414
Db |||||
Qy 1047 GACCTGAAA 1055
Db |||||
Qy 415 GAATGAAA 423
Db |||||

RESULT 13

US-09-221-017B-960/c
; Sequence 960, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 960:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3250 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...3250
US-09-221-017B-960

Query Match 4.8%; Score 50.6; DB 3; Length 3250;
Best Local Similarity 49.1%; Pred. No. 2.2e-05;
Matches 134; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy	783	TATTAAGAAAACACGAAAGAAACAATCTATATACATTCATTATGTAAAGAGACACGGAAT	842
Db	1540	TATCAGGGGAAGAAAGAGGGACTTATTTTGGATTACACAACATGATCCGAGAAGATGCCCAT	1481
Qy	843	CGCGCTGACGGCTTTTCGACAAGGGGAAGAAAAAATGCTGTTTACGAAAAATGCTGAATGTTT	902
Db	1480	TTATTGTATGGCTTTTCCAGGAAGAGAGGGTACGCTCTTCGGCCAACTCAATCTGTCT	1421
Qy	903	ACGGGGATCGGCCCAAAAGGAGCGCTTTCGATTCCTCGCTTCGGCGGATCCGGAGACGGGTG	962
Db	1420	AGCGGTGTCGGGGCTACGACGGCACAGCTCATCTCTTCTCTATGCTCCTCAAGAGCTG	1361
Qy	963	ATTGAAGCGATCGAGATGAGGACGAAGCATTTCTCGTCAAAATTTCCGCGCTAGGCAAA	1022
Db	1360	GCCGCATCTATACCACAGGCAGGCCGATGCGCTGAAGCGGTGAAGGCGATCGGTCGTG	1301
Qy	1023	AAAACGGCAAGGCAGATCATCTTTGACCTGAAA	1055
Db	1300	AAGACCCCTCAGCGTATCATCTGTGGATCTGAAA	1268

RESULT 14
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent NO. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses thereof

Query Match 4.6%; Score 49; DB 3; Length 1830121;
Best Local Similarity 49.4%; Pred. NO. 0.0036;
Matches 127; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 799 AAGAAACAATCTATACATACCATTATGTAAGAGAAGACACGAATCGCTGTACGCTTTT 858
344955 AAACAACTTTTATTCACCCATCTTGTTGTCGGAAGATGCTCACTTATCTTTTGATTG 344896

Qy	859	CGACAAGGGAGAAAAAATGCTGTTTACGAAATGCTGAAATGTTACGGGGATCGGCCCAA	918
Db	344895	CCCAAGAAAACAGACCGCACCTTATTTCTGTGAATTAATTTAAAAACAAATCGGGTGGGGCCTA	344836
Qy	919	AAGGAGCGCTTGCAGATCCTCGCTTCGGCGATCCGGGACCGGTGATTCGAAGCGATCGAGA	978
Db	344835	AATTAGCGCTTAGCGAATTTATCCGCCAATGTCGGTCGAACAATTTGCGCTATGCGAATAGAGA	344776
Qy	979	ATGAGGACGAAGCAATTTCTCGTCAAAATTTCCCGCGGTAGGCAGAAAAAACCGCAAGGCAGA	1038
Db	344775	GAGBAGAACTTCTAAACTTACTATAAATTCACGCTGTGGCAGAAAAAACAGCTGAACGTT	344716
Qy	1039	TCATCTCTTGACCTGAAA	1055
Db	344715	TGTTAGTTGAGCTCAAA	344699

RESULT 15
 US-09-643-990A-1/c
 ; Sequence 1, Application US/09643990A
 ; Patent No. 6528289
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert D. Fleischmann
 ; Mark D. Adams
 ; Owen White
 ; Hamilton O. Smith
 ; J. Craig Venter
 ;
 ; TITLE OF INVENTION: The Nucleotide sequence of
 ; the Haemophilus influenzae Rd Genome, Fragments
 ; thereof, and Uses Thereof
 ;

Query Match 4.6%; Score 49; DB 3; Length 1830121;
Best Local Similarity 49.4%; Pred. No. 0.0036;
Matches 127; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

```

Qy 799 AAGAAACAATCTAVACATACCATTATGTAAAGAGAGACACGAATGCGCTGTACGGCTTTT 858
Db 344955 AAACAACCTTTATTCACCCCATCTTGTGTTTCGGGAAGATGCTCACTTACTCTTTGGATTGG 344896

Qy 859 CGACAAGGGAGAGAAAAAATGCTGTTTACGAAAAATGCTGAATGTTACGGGGATCGGCCCAA 918
Db 344895 CCCAGAAAACAGACCGCACCTTTATTTTCGTGAATTAAATTAACAATAATGGGGTGGGGCCTA 344836

Qy 919 AAGGAGCGCTTCCGATCCTCGCTTCCGGGATCCGGGAGCGGTGATTGAAGCGATCGAGA 978
Db 344835 AATTAGCCTTAGCGGATTTTATCCGCCATGTCCGGTCGAACAATTTGCCCTATGCAATAGAGA 344776

Qy 979 ATGAGGACGAAGCATTTCTCGTCAAAATTTCCGGGGTAGGCAAAAAACGGCAAGGCAGA 1038
Db 344775 GAGAAGAATTTCTAATACTTACTAAAAATCCAGGTGTTGGCAAAAAACAGCTGACGTT 344716

Qy 1039 TCATCCTTGACCTGAAA 1055
Db 344715 TGTTAGTTGAGCTCAAA 344699

```

Search completed: February 27, 2006, 06:36:34
 Job time : 245 secs